

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTC
AGTGACTGAGGCCAGGCCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCCTCGCCGCTGTG
CCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAGGTC
TTAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTTCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACCTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGA
 TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
 CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
 ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
 ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
 ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
 CAAGGTCATTACCAATGTGTGGCTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTT
 AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
 GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
 TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
 TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGGGGATTCCACG
 CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
 CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
 GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAAGTCA
 CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCCAGATGAAGAATTCATGGTTC
 TGAAAATAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
 TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
 CACAAACACTACAATCTGAGCACCTTCCCTGATTTCAAATTATTTAGTGTAACAAACG
 GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
 ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
 CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGAAACATGGC
 TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
 GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
 GCAGACCTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
 GTCACCCTTCAGGTTACACACCCTGCCTCGTCTTTATTTACATTCTAGAAACAA
 CAATACAGGTAAGTTCCTTGTCCAGGCAGTGGGAAGAGAGAGGACAATTTGGA
 GCAGAGGGAGAGTTAAAAACCCACATAGAACCCCAAGTCTCCGACGGCATAGACAC
 AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
 AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGGCTCACAG
 TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAATA
 ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
 TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
 TCCGGAGGGGAAAGTAACCACGTGATTCTTACGGAAGCAAGCATGACTTCTGCTCC
 AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
 TGGGACCATCCAACTGGGAAAGATTGAGTGGAAACAACACCACTTCCCAGCCCCC
 TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACCTCCCTTG
 CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
 ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAATAGCTCCTCTTTTACC
 CACAGGTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCACCAGC
 TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCTTGGCTTTTCAGTCATTCCCCACGACAAGATGATGGTGGAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCCTCCTCAGC
AGTGGGGCGCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
TGGTCAAAACCCAACCTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAATAAAGTGCAAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACCTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTCCAAGTTGACTGCA

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCAATTTCAATTCTCCAATGACAGGTTACATTTT
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAAGCC
ACAGCCATTTCGGCACTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCTCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTG
GCCAGGCAGTTATTAGCGTGTCACTGATGGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTGAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACCTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGAA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACCTCTTAACCTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCTCGTGTCTGTTTCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGCGGTCTGTGTAAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

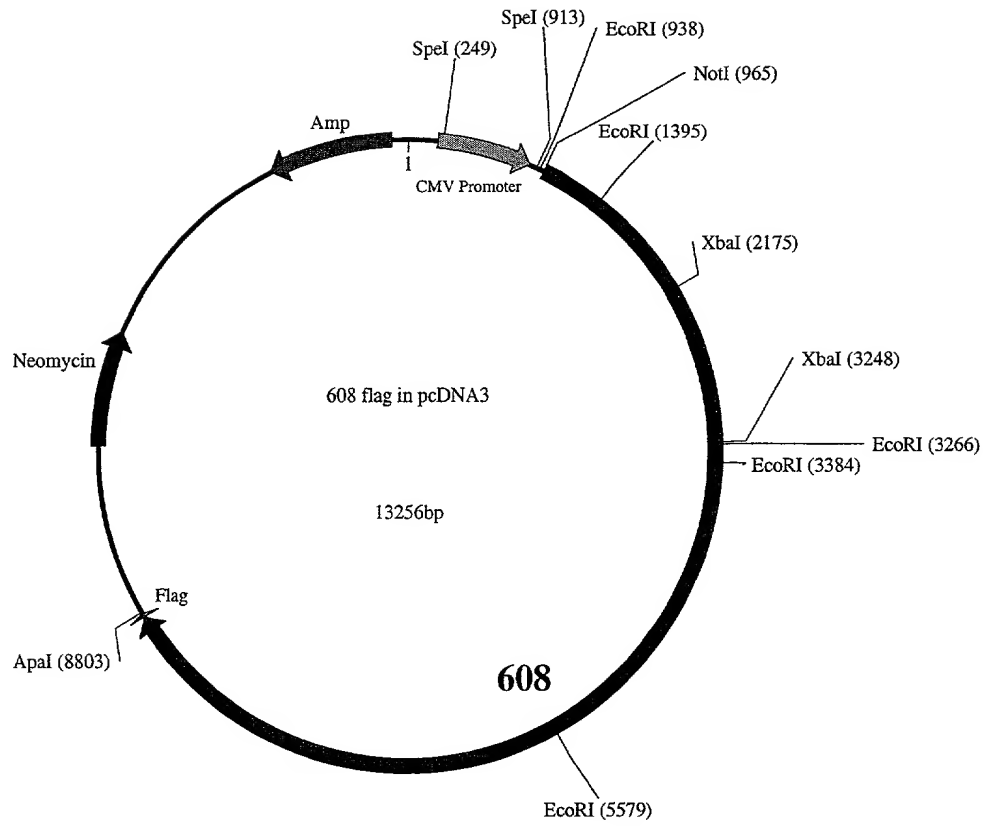


Figure 3

MQVRGREVSGLLISLTA VCLVVT PGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
 RINLGYNLSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
 RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
 FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
 IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
 ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
 VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADV
 ADPFWFQQEKIVLQLNRTATTSLTLQIQFSTDAQIALPRAEMRAERLKWMTMILMMNNPK
 LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
 DSFDAGLYHCISTNDADADVLT YRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
 PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSFVKV
 SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKL SASALTGSEAGKQVSGVHRK
 NKHRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
 KPVPLAVPLVELTDEEKDASGMIPPDEEFMV LKTKASGVPGRSPTADSGPVNHGFMSTI
 ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
 IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
 SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
 KNPHRTPVLRHRHRTVRPAIKGPANKNV SQVPATEYPGMCHTCPSAEGLTVATAALS
 VPSSSHSALPKTNNVGVI AEESTTVVKKPLLLFKDKQNV D I E I T T T T K Y S G G E S N H V I P T E
 ASMTSAPTSVSLGKSPVDNSGHL SMPGTIQTGKDSVETTP LPSPLSTPSIPTSTKFSKRKTP
 LHQIFVNNQKKEGMLKNPYQFGLQKNPA AKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
 TMAATQNKGT EVVSGARSLSAGKKQPFTNSSPVLPTISKRSNTL NFLSTETPTVTSPTAT
 ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALT TADTPLAFSHSP
 RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQUESTTVKRASDTP
 PPLLSSGAPPVPTSPPPFTKG VVTDSKVTSAFQMTSNRVVTIYESSRHNTDLQQPSAEAS
 PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSWPPEHKYQLKSYSETIEKGK RPA
 VSMSPHLSLPEASTHASHWNTQKHA EKSVFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
 GKAASFTVPANS DVFLPCEAVGDPLPIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV
 SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
 RPTVSWILANQTVVSETAKGSRKVWVTPDGT LIIYNLSLYDRGFYKCVASNPSGQDSLL
 VKIQVITAPPVII EQRQAIVGV LGGSLKL PCTAKGTPQPSVHWVLYDGT ELKPLQLTHS
 RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVILTVEEGETIPRIETASQKWTEVN
 LGEKLLLNC SATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
 YLCVARNKMGGDDLVMHVRRLR LTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
 SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFN NVGMAEEGDYICSAQNTLGK
 DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
 NDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
 VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLILYKATRKNKSGKYRCAARNKVGYIEKLIL
LEIGQKPVILTYEPMVKSVSGEPLSLHCVSDGIPKPNVKWTTGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVA YPPRIINYLPNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
22

Figure 4

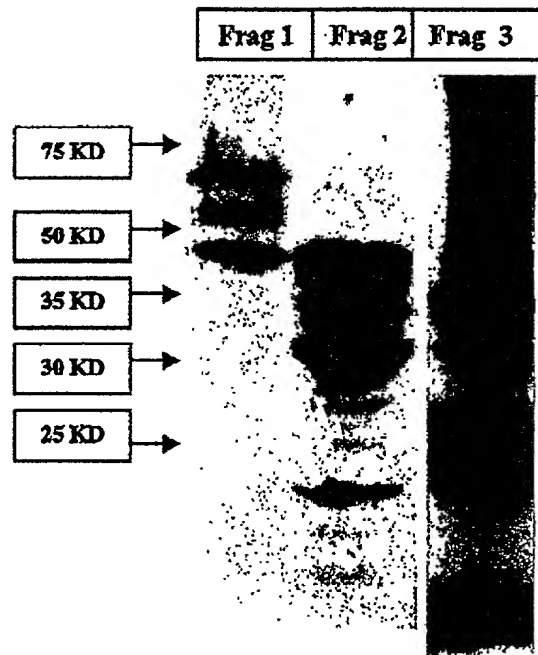


Figure 5

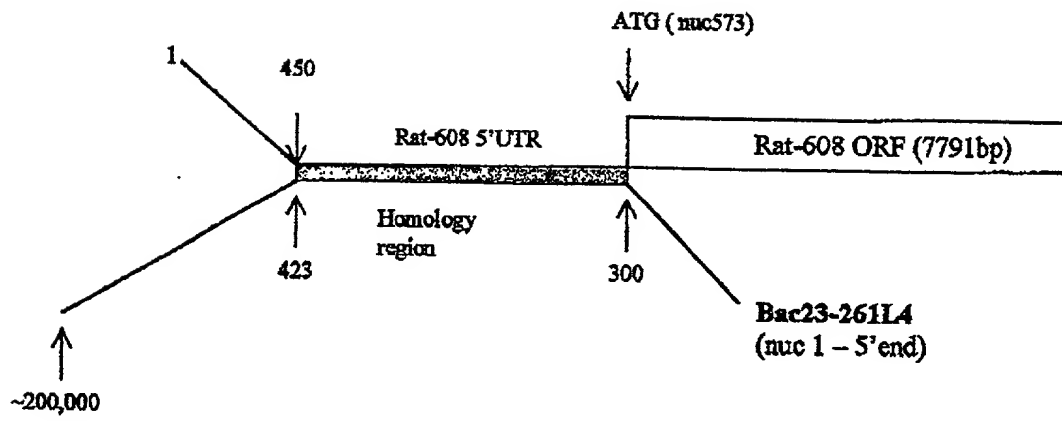


Figure 6

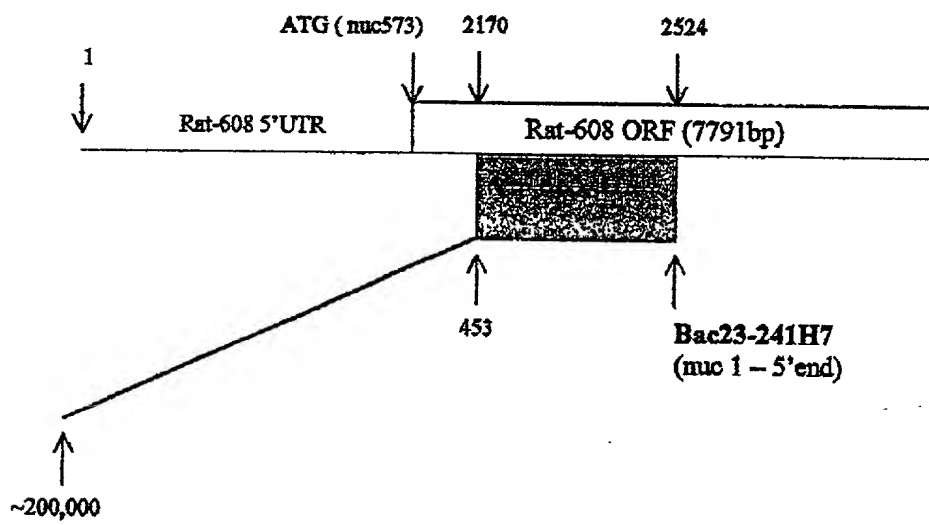


Figure 7

Nuc 1

↓
 TTTTGGAAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAATGTGGTCCATTTA
 TCCAATGGAACTACTCAACTTATTAATAAACCAACGACTTTCATAAAATTTTAGGCAAAT
 GNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACACTCATGGTATGC
 ACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTAAAAAGGGAAGAAGACACATAGC
 TTTTGTGTGTATAATTAAGATGGAAATTTGCCAGTGTCTTTGGCTTATGAGTGAATCT
 TGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAGTGTAGGAAGAAGTAGTTAATCA

Nuc 390

AGGTGACAAAGTATCCTGACACATTGGGAAAAGACCACAGTCCAGGAAAAGTGAAGTCTTAA
 GGATTTCATATTAACCTCCAGTTCCTCCATGTGCAGCTCTGAGACTTTGGCAGATCAGACACTT
 AACTTCAACAGCTTCTACACAGAGCAGTTACTATCCTTGCACCTTACAGATGGAGTGTGA
 CCATTAAGCTGCACCTGAAACATGAGTCTGACTTGTTAATAATCTTAAAAATACAAATTGTGT
 TGTAAAGTATGTGACCAAGAGCATGGTTCATGCTATTAACCTTTGATGTTCTATGGACTCT
 TAATTTTATGGTAGAAATGTCAACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAG
 GATGGCTTTCAAGTTTGAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCA
 ATTAAACCCCAAGTAGGCATAATTTAATGCTTACTTCATCAGAATATGTCTAATTGAC
 TCTTCTAAAAAGACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTACTGACATAAT
 AAATAGGAGAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTTCTAATAGGTCTT
 ATAGGTATATATAAGAGGTGGGAAAAGGAATATTTGAAACATCTAGAAGTAAAATGATCCTG
 AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTGTGTTGCT
 GGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATACAGAGCTAT
 GACCTTAGCTGGAAGGAGCTTTGGAAAGAAATGACAAGCTGTAGCTGCCAGAACATCTCTG
 GATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAAGAAATATGGTTGTCTCTG
 CAGTGTCTAAGGACCTTGTCTGTGCCACACAGGTCTCCAGGCTTAGTGCTAAGCTGTGACTGC
 TTGACTCTTTAATCTACCTTGATCAATTAAGTACTAGAAATCAGTTGGTGATTAGCAACTG
 GATATGGAATATTACTAATTTGTACCCAAAGCCAGGCCACCTCAGCTTTGGCAGCTCCATTC
 ATTCTGTGGAGCCAGTCACGTGGGTTTGAATCAACTGTACTGTTTCTACTTACAAGACGC
 ATTACCTGAGATGAGTCATTTTCTTCAAGTCTTTTGAAGAGTCAATTAGACATATT
 CTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGA
 TGCTATTATAATGTGGAAATAAAGTACTTTAAAAAAGGCTTATGATACTTGTCTACA
 GAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTATTTTCAATCCATCAATTAA
 AAACCTTTAGAAATTTTGTAAACACAAAGATCCCTAGGCCCCCTGCCCTAGGATGGTCTGTA
 TGGTGGCCCTGAGAGATGGAGCTTAAGAACTTACTTGTCTCCAGGAGCACATCTTCAGAAC
 ATCTGCCCTCAAAACATTTATCCAAATGCTCATCAAGGCTCACTCACATGTGCTTCAACC
 ACAGGGATTAAACAGTCAATTTTAGTCACATTTCTCAAAACGGTGGAAAGCCTGACAGGAA
 CAGGATGTATCAGGATAACATCCAACTTACAAAAGGATGTCTATAACCTCACCACAACA
 AACAAACAAGCAACAAACCCATAAAAATTATCAAGGCAATGAACATAAGCCATATGCA
 GAAAAAGTATTATATGTTCTCATTTGTGGGGTGTCTTCTTAATAGTCAAAATATGCAGAAT
 ATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTTGTAGATTAGAGGACA
 CAAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAATGTATAATATGAGGACTGTA
 TTTAATAATCTATTAAAGATACACAGCAAAACGAGTGTATCTTACTAACACACACACTTAC
 ATAGAGAGAAATAAAGTGATAGATACGTTTGTCTTATCTTCATGTAGCTGATAATTTCATAT
 TGTACACCTCAAACATAGATAACCAACAAAGAGGAAGAGGATAGGTGCCCTCTCCAGGGC
 GGAAGAGTACATTGCAAGTCAAGACCATTTGTGTAGATGTACCACATGGAGGAGCTAGA
 GAAAGTAGCCAAGGAGCTTAAAGGGATCTGCAACCCTATAGGTGGAAACAACATTATGAGCT
 AACCAGTACCCCGAGCTCTTGACTCTAGCTGCATATATCAAAAGATGGCTAATCGG
 CCATCACTGGAAAGAGAGGGCCATTGGAGCTTGCAAACTTATATGCCCCAGTACAGGGGA
 ATACCAGGGGCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGTGGGTGGATATGGG
 GGACITTTGGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGAAAAA

Nuc 2317

AAAAAAGGAAAGGTCAGACACCTCACTTCACTGCTATCTCAACTT
 GCAACAGAAGGGGAGTCACAAACCCAGGACAAACACAGTGATTGAAGCGTCTTTGAA
 TGTATTGCTGTTGTTGTACCACCATCATTAGCATATATTCATTGTGAAAACTTACGGGGT
 CTATGACATGTTTTTTTATCAAGTATATCATATGCTGTGACATATTTGGCACCCTACCA

GCGCCAGGCCCTTTGCCCGCGCCCCAACACACACACACACACACACACACAC
 CACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCATCATCTGCTCACTCAGC
 CACCCAAGCTTAATCCTTTTCTTTCCTGCAATAGTAOCTCTOCTATTTTATGTCTAGGTT
 CCCCCTCCCCCTGTTAGGAGATGGGAGAGGTACGAAAGGAAAGAATTTGTAGCCCCCTGA
 GCGAGCGCGGGCCACAGAGCCTGCCACCAGACAGGAAAAGGCCAGGGCTTACCAGCACA
 GGAGGAGCAAACCTCGCAGGCGAGCCTGGGTTGGCGCTGGTGCTCCCGGGTCCGATGGCCCCG
 CCCATTCCAGAAAGCCGAGGCTATAGCTCGCTCAOCTGCCCGGCCCTCCTCCGAGTGAAG
 ACCCTTAGAGGCTGACGACACCCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGT
 GAGCGGCTGCTAATCCCCCTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGA

GCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAAGGCTCCAGGAAAAAGGAAAGT
ACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAAACATCGCTTCAGTGA CTGAAGCCA
GGCAAAAAGGACGGGAAGGATTATATGTAGCCTGGGAAGCTTTCATAAACACTGTACGTGT
TTGTGCAAAAGCAAGCAATTTGAGGAGAAAACGCTGGGACCTCGGAAAGGAGAGTGATGAGT
AGTACTTTGTAAGTTTAGGTTGAGTTTGAGATACTAACTAACTACTATTGAGGGGAGGAAGAAG
AGCATTCCAGCAGCAGCAGCAGCAGCAGCAATCAGATAAAGGAAAAGCTTTGGTTAGTTTGGAA
ATGTATGATAACCTATAAAATAACAGAAAGCGCTCCAGTTCTCTGAAGAGTCA GTCCCCAGCTA
GTGAAGACTAAGCCTACTAAGCCTTTGTCTCCCGTTGGAAGCAAGAAAGTTCTTTCATCAGG
TGAAGAGCTCTCTCAGAAAGATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGAC
AGAAGAGCTCAGGTATTGCCAACTCTTT

TGTTAAATACAGTTTTGAGGCTTAAGTGTACGGGAACTCATGTGGTATTTCATTTACGGGCTCTCTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTCTGTTTCCAGCTACCTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAATCATCACGCAGAGAGTATTCAAAAAATTTTCAGTATGTAAACAGTACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTGGGTTAAACTCATTCATAACAGCTCCAGGTGGGAGGGATCACTGAGGCTTCGCCAGCTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCCTCTCTGGCCATGCTCCCATCACTGTGTAGTAAGCAGAGGGGGTGTTCCAAAGCAGAGAAAAGAGCAGACAGTGT/TGCCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACTGTCTCTCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAA/GAAATTTGACCTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTATTTTGAAGGTAAAGCCAAATAAAATAAACACATTAATAATAACAATTGACAAAATTGATCATATAAAGAA/TGATCTCATAAATGCAATGTGAAAAATATATAATTTTATCTATTTACTGGTGCAAAGTTTCTAAATTGCACATGTACTATTTTTATATTATAAAAAATTTTTTAAATGTATATAAAGTGTAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACTTTAATTTGTCCGAT/AACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGAAATGAGAAGCTATG/AAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCAATGGGAAGCTAGCACACACTGGTTTATAAAGGAACAGTTTCTGTAGTCTATGCAAGTTTACTCTAGGGAATAAGAGTTTCTCTTCCAGATTTCCTAGCATTTTGTGTGCATCATTTATCTCTGTGATGATGAGCATTATAAGTGAATAAAGATAGGATCTCAAAGGAATGTCAATTTGGATGCCCTGAACAAATCTTCAGGCTCTTCTTCAGTTCACTAGTCTATTCATTTATTGGATAATTGGGGGATGGTGT/TAATTTTTTGCAGTCTTATGGAATTTCAAAAAACAAAAACAAACAAACAAAAACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACA/GAATTTATTGCTACATTGTTTATTATTTCACGACAGCCAAGAATGTGGAACCAACTTACGTA/GCCGTAAAAATAGAAAGGATAGAAAGATAATGTTGAAATGTGTACAACAGAGTCCCATGTG/GCCCTAAAAGATGAAAGCATAGACATATGCAAGAAATGGATGCAACTGGAAATCAATTTGGCTAATCAAAACAAGACAGAGCTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACA/GAAGCTAGATTACACTTGTACGTGCGCATGTGTGTTAGAATTTTATTAGTTATACACT/ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAACATAGGTG/CCAGGATGAAATGCATTAATTTGTATGCTAACTAAACACAGACAGGAGGCCACAGTTCA/AACCAGGGTGAAATCCCAAGCAGATAGAGAGGGGAAGTAGACACAAAGTTTCCGACCTAAC/CAAGAAGCCATTTCAGATTGCTGCCTGCTGGGAGGGGGGTTCCAGTTTCTCCAGCTGAC/ACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGGTTTTTGTGCTATT/TTTCATTTTTTTCTTACGTTTTTGTGTGTGGTGGTGTGTGTGGTGGTGGCTGTGGTTTTCT/ATTTGTTCTTTTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGAAGCTGGAAACG/ATCTGGAAGAAGTTGGGGAGAGAAAGTGTATGGAGCATTTTAAACAAACAAACA/ACAAACAAAGGTTTCATTTTGCCACAAAAGGTTGTAATTAATTAACGATACGACT/CTTAAAGAAAATATTCCCAATTATTCCAGAGTTGCTATGTATGCTGTGCCCTAGGACTTTTG

▼ Exon 2 (Nuc 6559)

CTTGAACCTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGGATGCAAGAGAGGGCAGGGAA
 GTCAGCTGCTTGCCTGATCTCCCTCACTGCCATCTGCCCTGGTGGTCACCCCTGGGAGCAGGGTC
 TGTCTCTCGCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTGGGAACCTGAAC
 TCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

↑ Nuc 6755

GGACCTTGCCTGATCTCCTTCTCAGAGAGGGGACCACTGATTTTCCTGGTACTTTGCCCCC
 AAACACCTGTGATTACTTTTAATAGTTTCTTCTAAATGGGTTCATACAAACCTTATATTG
 TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAAGTTGAAGCCCCTCTTAGTT
 GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTTAAATGAAAAAGAT
 ACAGTTTCTGTAAACAATCATTATGATACCAAGGAAGAGGAATTGTCAATTGAATATTTTAA
 AAAAGCATTCTTTTGCATTTTATAAATACCCATTACAAAATGGCTTACTTAAAATACTTG
 CCTTACTAAATCTGACAAATTATGGTGATATTTGAAGGTTTATGAAAAATTGTTTATGTGT
 ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT
 ATGCCAAACACTATTGTTTACCGTTTGGAGGAAGTAATGGGGTGGGGGAGCAACAAGGGT
 TATAACCGTATACCCAGTGCCTTGAAGCGATTGCAAAACAGTAAAGACTGACATTGTGTT
 CTCCCTATGAGGGAGGGGCTTGGGCTGAGCATTGTGCAATGAGCATTGTCTCATTGTCT
 GGCAGGTTTATGATAACTTGACCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACT
 GAGAACATGCCTGAAGAAGATCAGATTATAGGCAGGCTGTGGGGCATTCTTAAATTAG
 TGATTCATGGGGCAGGGCCAGTCCATTGTTCTGTGGTACCATTCTCAGGCACTATTAAAA
 AAAAAAACAAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGA
 TCTCTGCATCAGCTCCTGCCTCCAGGTTCTACCTATTGAGTTCTGTCTAGCTCCCTA
 CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTTGCTG
 TTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGTATAACA
 CCTTGGACTAAATCTGTATCTATAGCTGAGGAAAAATGGAGCATAGAAAGTCTCCAGA
 CTACACCAGAGTGAATCTGGGCTGAGCTTAGAATCACACCCACGTGCACTCCACTGCC
 GGGGCTTCTTAAACCGGAACACAGTTGTAAAAAGGGAATTTCTGTTTGTTCCTTTTGACA

Exon 3 (Nuc 8089) ▼

TGTGGACTTTAATTGACGATTCTCTGAAGCTGAAAATGATTTTTCAGGTATAACAGC
 CTCCTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATGCTGCACA
 GCAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCTTGACAG

Nuc 8218 ▲

GTGAGATAGGTAGAGGGTGTGAGGGCTGAGAAGAGAGGTGCAACTGTGGGTATACCC
 AAAAGCTGCTGATTCCCGTGGGAGACATTCTATAAGCATTCTATAAACTAGAGGCAGATA
 TCAAGGAAGGATTCAATTGTAATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATG
 CTGGGGAATAATGCTTACACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAA
 AAAGCACAGCATGGCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTTCAAGT
 CCCAGAGACTCATTTGGCCAAACACTTATTCAATCAATGAAGTCCATATTCAATGACAAA
 ACTTGACTCAGAAAATAATGTGGAAGCATCAGGAAGACAGCCAACATCTGGTCTTACT
 CATGCTATGAATAAGGGATCCAGAGAGAAAGGAAGAAAAGGAAGGAAGGAAGGAAGG
 AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
 AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
 GGAGAAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
 TTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
 AAGGGCAGTCTTTAACAATAGTAGCATTTTATAAACCATTACAAATTTTGAGTTTCTCTAC
 TTTTATCTCTACCATCTTCAAACCTGAAACTACAATTATCCACAAAATGAAGAAAATGC
 TGTAAGAGTTTTCACACACCGAAGTGGGAACTTAAGGATTAGACAAGTCTAACAATGAG
 AATGGGGAGAACAAAAGAGACTGCACAGGGAGCCCTTTCTCTGCTTATAATCTTGACAC
 TTGAGAAGCTAATTGACGCTGCATGACTACTCAACTCTTTAAGCAAAACATGCTGTTGTT
 ATGAAAAGCACAATAAAGTACATATGTCCATAATATTTCATCAAAATTTGCATGCAGCAC
 ATAATAGCAATCAAAGCAATAACACCCACTGTTTACAGAGACTTTAAACATGAAACTGGA
 ACTATGTCTAGTGTTTGACTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTG
 ATTTAGGTCATCAGACAGCATTGGAACATGTATCTTCAGGAGGAATCATTATGTATCCT
 GCATGAAATTTCTCCACCTATGTTTATTCTTATAGCCAGGTTTTCTCTGATGGAGAAACATT
 GGGTTTGAGGTTTACTCCAGGTAACATTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG
 CTTTTATTTATGAGGGATGTTGGTATTCCAGAAAATTTCTCTTTGAAGAGATTACAATTTA
 GGTCAAAACAGAAAATATGTAAAAAGTTATTGTTTTATTAGTATTTCATGTTCTTTCTT
 TTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGAAAATAATC
 AGAGAGGGATTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTTTTGCTTCTAATT
 CAGAAATCAATTAATTCATATTACTATAAAAGACAGCAGCCAGATGTGTGCCAGCTGAG
 GAGTGATAAACTGTGTAACGTGAGTGCTATGTAGAAACAGAAAGGAGTGAAGGGTTGA

TGTGCGCTGCAACATCTTGAAAACATTCGGCTACATGATGGAAGCCAGGCACAAAAAGCC
 ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAAATACATGGATTCTTAGCAAAACAGA
 GTAAGATGTTACTTAGGGTCAGGAAAAGATTAAAAAATACTATTGATGTGGAATG
 ATCTTAATTTGGGGAAAAGACAATTTCTAAGACGAAATAGTTGAGGTAGATATAGTTAT
 ATCCCTGTGGATATTGTAAATAAACAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGG
 GCAGGAGGAAAGTGAAATGAGATGGTAGAAAGGAAAGTCATATAACCATGGCTTCTCTCGTG
 GGTGGAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATC
 AAAACCAACAGGAGTGAGGGAGACAATAGGAACCAATGAGAGGCCAAAGTTCATGGTCAA
 TGTGTGTGGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATC
 TAAAAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGGG
 GAAGTTGAAGTTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTACGTTT
 TTTATTATTAATTACTCTTGCTGTCTATTATTCATCATCATTATCGTCATCATCACTACTA
 ATGCTTTTACCATAACACAAATGTAGGCAGAAGAGTGTAATCCACTTAGTGAGGCAATCTT
 GGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAGGACAGTGAGAGGGGAA
 ATGAACAAGAAAAATGTGGACACATGCACAAAAATTCATAGTCCACTACATTACTTTG
 TATTCTAATATTAAGAAAAATAATAAACCCATTCTGTGCACCTATCACCCAGGCTCAACAG
 TTATCTTGGCCACAGATCCTGTCTCACTGCATCCTGTCCACCTGAGTCCACTTAGCGTTCTG
 AATCCAATCCAGGGCATGATGCTTACTCCTACACAGAACTAAAGATTAAAGAGAGTTTAA
 AAGTAACCATGACATCTCTCTGTTCTTTAGCGATAAGTTCTTAATATTTATGGCTGCTTGT
 GTATGTTCTAATTTCTCTAATATTGTGCATTTAGTTGGCAACTACTTTGTTTGAATTGAGT
 TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTTCTATATTTATAAACCTTTCTCTCT

Exon 4 ? (Nuc 11286) ↓

TTGTGAAAGTTCCTTTGAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATG
 AGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGG
 TTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTTGC
 TCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACATTGTCTCTTT
 GAGCTATCTCCAGATATTTAAAACTCCTTCATTAAGNACCTGTACTGTATGATAACTTCATTG
 AACTCCCTCCAAAAAGAAATGGTCTCCTCTATGCCAAACCTAGAAAGCCTTTACTTGCATGGAA
 ACCCATGGACCTGTGACTGCCATTTAAAGTGTTGTCCGAGTGGATGCAGGGAAACCCAGGTA

Nuc 11680 ↑

ACTATCTTGTGTTTGTGTTCTTTTTTATARKACGTAATTTCTCAATTTCAATTTAGAATGA
 TATCCCAAAAGTCCCCATAACCTCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGG
 CCTGGCATTCCCCCTGTACTGGGGCATATAAAGTTTGCCTGTCCAATGGACCTCTCTTTCC
 AGTGATGGCCAACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG
 TACTGGTTAGTTTCATAATGTTGTTGCACCTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

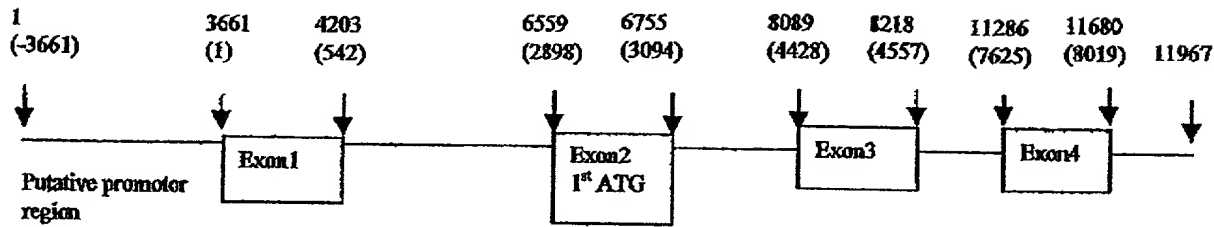


Figure 10

cDNA_rat	1	cgagagacgacagaagggttacggctgcgagaagacgacagaagggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgcctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgcactgaggccaggcaaaacgcgcgggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggaacctttcatagacactgatgacacgtttacgcaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaacgcctgggccttoggaaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	cttgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaatggatgatagcattaaaataacagaagcgccctccagg
genomic_hu	1	-----
cDNA_rat	401	tctctgaagottcagtcctccagctgaaagccagaaaagactaagccac
genomic_hu	1	-----
cDNA_rat	451	taagccttttgatccctttggaagcaagaactttccttcctggggtga
genomic_hu	1	-----
cDNA_rat	501	agactctctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcagggtgagaggcagggaagtacg
genomic_hu	1	-----
cDNA_rat	601	cggottgttgatctccctcactgotgtgtgcctgggtcaccctggga
genomic_hu	1	-----
cDNA_rat	651	gcagggcctgtcctgcgcgtgtgcctgctatgtgccacagaggtgcac
genomic_hu	1	-----
cDNA_rat	701	tgtacatttcggtaacctgacctccatcccagatggcatcccgccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaatttaggatataacagccttactagattgacagaaaacg
genomic_hu	1	-----
cDNA_rat	801	actttgatggcctgagcaaacctggagttactcatgtgcacagtaattggc
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgagtccttgacaggt
genomic_hu	1	-----
cDNA_rat	901	cttaaaaatgagctataacaaagtccaaatcattcgggaaggatactttct
genomic_hu	1	-----

cDNA_rat	951	acggactcgggagcttggtccggttgacactggatcacacaacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaacctgaggccttttatggacttacctcgctccgcttggtaca
genomic_hu	1	-----
cDNA_rat	1051	tttagaaggaaaccggctcacaaagctccatccagacacatttgctcat
genomic_hu	1	-----
cDNA_rat	1101	taagctatctccagatatttaaacctctttcattaagtacctgttcttg
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttcctgacctccctcccaaaagaaatggtctcctacatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacotagaaagcctgtatttgcatggaaacccatggacctgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	atttaaagtgggtgtotgagtggatgcagggaacccagatataataaaa
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaaagacagaagctcttcagtcctcagcaatgtcccctttgcat
genomic_hu	1	-----
cDNA_rat	1351	gaaccccaggatctctaaaggcagaccctttgctatggtacocatgtggag
genomic_hu	1	-----
cDNA_rat	1401	ctttcctatgtacaaagccaaccattgatccatcactgaagtcaaagagc
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu	1	-----
cDNA_rat	1501	catagaacccctttggctccttgtctttgaacatgacanannntnctggaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatggtctgtagtatccaaaagccatcaaggacatcacca
genomic_hu	1	-----
cDNA_rat	1601	actgcattcactgaagaaaatgactacatcatgotaaatgcgtcattttc
genomic_hu	1	-----
cDNA_rat	1651	cacaaatcttgtgtgcagtgtagattataatcacatccagccagtgtggc
genomic_hu	1	-----
cDNA_rat	1701	aaottctggctttatacagtgactctcctctgatactagaaaggaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagcttaccgagactccttcactgtcttctagatataaacagggtggctct
genomic_hu	1	-----
cDNA_rat	1801	taggcctgaagacattttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgagctgaacagaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagcacattacagatocagttttccactgatgtcctaaatcgc
genomic_hu	1	-----
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaattggaccatgatcc
genomic_hu	1	-----

cDNA_rat	2001	tgatgatgaacaatccaaaactggaaogcaactgtcctggttggcggcact
genomic_hu	1	-----
cDNA_rat	2051	attgccctgagctgtccaggcaaaggcgacccttcacctcacttggaatg
genomic_hu	1	-----
cDNA_rat	2101	gottotagctgatgggagtaaagtgagagccccttacgttagogaggatg
genomic_hu	1	-----
cDNA_rat	2151	ggcgaatcctaataagacaaaaatgggaagttggaactgcagatggctgac
genomic_hu	1	-----
cDNA_rat	2201	agcttttgatgcaggtotttaccactgcataagcaccaatgatgcagatgc
genomic_hu	1	-----
cDNA_rat	2251	ggatgtttcacatacaggataactgtggttagagccctatggagaaagca
genomic_hu	1	-----
cDNA_rat	2301	cacatgacagtggagtccagcacacagtggttacgggtgagacgctcgac
genomic_hu	1	-----
cDNA_rat	2351	cttccatgcctttccacgggtgttcagatgcttotattagctggattct
genomic_hu	1	-----
cDNA_rat	2401	tccagggaaactgtgttctctcagccatcaagagacaggcaaattctta
genomic_hu	1	-----
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgcacaaagatcaaggtcat
genomic_hu	1	-----
cDNA_rat	2501	taccaatgtgtggctgccaaccocatcaggggcccacttttccagttttaa
genomic_hu	1	-----
cDNA_rat	2551	agtttcagttcaaaagaaaggccaaaggatgggttagcatgacagggagg
genomic_hu	1	-----
cDNA_rat	2601	caggtggatctggacttgagaaacccactccagtgtttcccttaagcag
genomic_hu	1	-----
cDNA_rat	2651	ccagcatctttgaaactctctgcatcagctttgacagggtoagaggctgg
genomic_hu	1	-----
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaaacatagagacttaatac
genomic_hu	1	-----
cDNA_rat	2751	atcggcggtgtgggattccacgctccggcgattcaggagcataggagg
genomic_hu	1	-----
cDNA_rat	2801	cagctccctctctctgctcggagaattgaccgcacacgctgggcagcact
genomic_hu	1	-----
cDNA_rat	2851	tctagaaaaagccaaaaagaattctgtgccaaaaagcaagaaaatacca
genomic_hu	1	-----
cDNA_rat	2901	cagtaaagccagtgccactggtgttccctcgtggaactcactgaogag
genomic_hu	1	-----
cDNA_rat	2951	gaaaaggatgcctctggcatgattcctccagatgaagaattcatggttct
genomic_hu	1	-----

cDNA_rat	3001	gaaaactaaggcttctggtgtccaggaaggtcaccaactgctgactctg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatcatggttttatgaacgagtatagcttctggcacagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tcaactgtgaatccacaaacactacaatctgagcaccttctgtatttcaa
genomic_hu	1	-----
cDNA_rat	3151	attattttagtgtaaacaaacgggtacagctgtgacaaagagtatgaacccat
genomic_hu	1	-----
cDNA_rat	3201	ccatagcaagcaaaatagaagatacaaccaacccaaaacccaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	tttccatcagtagctgaaattcgagattctgctcaggcagggaagagcattc
genomic_hu	1	-----
cDNA_rat	3301	ttcccaaagtgccacacctgtaacagggggaacatggctacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	ataccaacacatatagtagctttaccagcaaaagccagtagctcttgacag
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaacagaaagttatggacctcagatacctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacctagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctccagtcacccttcagggttcacacaccactgcctcgtctttattt
genomic_hu	1	-----
cDNA_rat	3551	cacattcctagaanaacaacatacaggttaacttcccttgccaggcaactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttggagcagagggagaggttaaaaaccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccagttctccgacggcatagacacaggactgtgaggccagcaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggacctgctaacaaaatgtgagccaagttccagccacagagtaccc
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgagaggggctcacagtggctactg
genomic_hu	1	-----
cDNA_rat	3801	cagcactgtcagttccaagttcatcccaagtgccctccccaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttggggtcatagcagaagagtctaccactgtggtcaagaaccaact
genomic_hu	1	-----
cDNA_rat	3901	gttactattttaaggacaaaacaaaatgtagatattgagataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatattccggaggggaaagtaaccaogtgattcctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgacttctgctccaacatctgtatccctggggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgctctgggaccatccaaactgggaaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaaacaacaccaacttcccagccccctcagcacacctcaataoca
genomic_hu	1	-----

cDNA_rat	4151	acaagcacaaaattctcagagaggaaaactccottgcaccagatctttgt
genomic_hu	1	-----
cDNA_rat	4201	aaataaccagaagaaggagggtgttaagaatccatatcaattcggtt
genomic_hu	1	-----
cDNA_rat	4251	tacaaaagaaccagccgcaagcttccaaaatagctcctcttttacc
genomic_hu	1	-----
cDNA_rat	4301	acaggtcagagttcccccagattctacaactctcttgacaagtcgoc
genomic_hu	1	-----
cDNA_rat	4351	accagctctgtctacaacaatggtgccactcagaacaagggcactgaag
genomic_hu	1	-----
cDNA_rat	4401	tagtatcaggtgccagaagtctctcagcaggaagaagcagcccttcacc
genomic_hu	1	-----
cDNA_rat	4451	aactcctctccagtgttctctagcaccataagcaagagatctaatacatt
genomic_hu	1	-----
cDNA_rat	4501	aaacttcttgtcaacggaaacccccacagtgacaagtcctactgtactg
genomic_hu	1	-----
cDNA_rat	4551	catctgtcattatgtctgaaacccaacgaacaagatccaaagaagcaaaa
genomic_hu	1	-----
cDNA_rat	4601	gaccaaataaaggggcctcggaagaacagaaacaacgcaaacaccacccc
genomic_hu	1	-----
cDNA_rat	4651	caggcaggtttctggctatagtgcatctcagctctaacaacagotgata
genomic_hu	1	-----
cDNA_rat	4701	ccccottggctttcagtcattccccacgacaagatgatgggtggaatgta
genomic_hu	1	-----
cDNA_rat	4751	agtgcagttgcttatcaactcaacaacotctcttctggccataactgaact
genomic_hu	1	-----
cDNA_rat	4801	gtttgagaagtaaccccagactttgggaataacaacagctttggaacaa
genomic_hu	1	-----
cDNA_rat	4851	cgttgttgagcaaatcacaggagagtaccacagtgaagagcctcagac
genomic_hu	1	-----
cDNA_rat	4901	acaccaccaccactcctcagcagtggggggccccagtgccactccttc
genomic_hu	1	-----
cDNA_rat	4951	cccacctccttttactaagggtgtggttacagacagcaagtcacatcag
genomic_hu	1	-----
cDNA_rat	5001	ctttccagatgaogtcaaatagagtggtcaccatatatgaatcttcaagg
genomic_hu	1	-----
cDNA_rat	5051	cacaatacagatctgcagcaacccctcagcagaggttagcccaatcctga
genomic_hu	1	-----
cDNA_rat	5101	gatcataactggaaccactgactctccctctaattctgtttccatccactt
genomic_hu	1	-----
cDNA_rat	5151	ctgtgccagcactaagggtagataaacacagaattctaatggaagccc
genomic_hu	1	-----
cDNA_rat	5201	tctccctggccagaacacaaaatatcagctcaagtcatactccgaaacct
genomic_hu	1	-----

cDNA_rat	5251	tgagaagggcaaaaggccagcagtaagcatgtccccccacctcagccttc
genomic_hu	1	-----
cDNA_rat	5301	cagaggccagcactcatgcctcacactggaatacacagaagcatgcagaa
genomic_hu	1	-----
cDNA_rat	5351	aagagtgtttttgataagaaacctgggtcaaaacccaacttccaaacatct
genomic_hu	1	-----
cDNA_rat	5401	gccttacgtctctctacctaagactctattgaaaaagccaagaataattg
genomic_hu	1	-----
cDNA_rat	5451	gagaaagggtgcaagctttacagttccagctaattcagaogtttttctt
genomic_hu	1	-----
cDNA_rat	5501	ccttgtaggggtgtttggagacccactgcccacatccactggaccagagt
genomic_hu	1	-----
cDNA_rat	5551	ttcatcagganttgaaatatccaagggacacagaaaagcgggttccacg
genomic_hu	1	-----
cDNA_rat	5601	tgottoccaatggcaccttgtccatccagagggtcagtattcaggaccgt
genomic_hu	1	-----
cDNA_rat	5651	ggacagtacctgtgctctgcatttaatacactgggcgtagaccattttca
genomic_hu	1	-----
cDNA_rat	5701	tgtctotttgtctgtgtgttttttaccggcaaggattttggacagacatg
genomic_hu	1	-----
cDNA_rat	5751	tcaaggagatcacagttcactttggaagtactgtggaactaaagtgcaga
genomic_hu	1	-----
cDNA_rat	5801	gtggaggggtatgccgaggcctacggtttcctggatacttgcaaaccaaac
genomic_hu	1	-----
cDNA_rat	5851	ggtggtctcagaaacggccaagggaagcagaaaggtctgggtaaacacctg
genomic_hu	1	-----
cDNA_rat	5901	atggaacattgatcatctataatctgagtccttatgatcgtgggtttttac
genomic_hu	1	-----
cDNA_rat	5951	aagtgtgtggccagcaaccatctggccaggattcactgtttggttaagat
genomic_hu	1	-----
cDNA_rat	6001	acaagtcatcacagctccccctgtcattatagagcaaaagaggcaagcca
genomic_hu	1	-----
cDNA_rat	6051	tcgttgggggttttaggtggaagtttgaaactgcctgcactgcaaaagga
genomic_hu	1	-----
cDNA_rat	6101	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu	1	-----
cDNA_rat	6151	accattgcagttgactcattccagatttttctgtatccaaatggaactc
genomic_hu	1	-----
cDNA_rat	6201	tgtatataagaagcatogctccttcagtgaggggcacttatgagtgcatt
genomic_hu	1	-----
cDNA_rat	6251	gccaccagctcctcaggctcagagagaagggtagtgttcttactgtgga
genomic_hu	1	-----

rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
.c_hu	890	CtatGGAAGCAGAATCAGTCCATAAAATGGAACCTTGGaattagga
rat	7401	acatccGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGttcggagcgag
.c_hu	940	atgtgaGGCTTTTCAGATTGAGCCGACTTTATCTGTGTGGcccgaaatgaa
rat	7451	ggaGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAG
.c_hu	990	ggtGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAATGCTGAG
rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAagtcategccccagctg
.c_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgccagctgg
rat	7551	gcaagcccgtagCACTGAAGTCTGTGTGGATGGGAACCCCCACCTGAA
.c_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTGTGATGGTAACCCACCACCTGAA
rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAcaca
.c_hu	1140	ataatCTGGATTTTACCAAATGGCACACGATTTTCCAATGGACCAcaaa
rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCTTTACAAagcaa
.c_hu	1190	ttatcagTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAaaca
rat	7701	ctcggaaacaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGC
.c_hu	1240	ctcgggaggtgCAGGAAAATATCGCTGTGCAGCTAGGAATAAGCTTGGC
rat	7751	TAcategagaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
.c_hu	1290	TAtattgagaaattagtcatttagaaATTGGCCAGAGCCAGTTATTCT
rat	7801	GACATACGAACCCAGGGatggtgaagagcgTCAGTGGGGAACCGTTATCAC
.c_hu	1340	TACCTATGCACCCAGGGAcagtaaaaaggoaTCAGTGGAGAATCTCTATCAC
rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAAGTGGACTACA
.c_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
rat	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAcT
.c_hu	1440	CCAAGTGGttATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATatT
rat	7951	GCATGAAAATGGCACGCTGCTCATCAAAGCAACAACAGCtcaagaccaAG
.c_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
rat	8001	GAAATTATATCTGTAGGGCTCAAACAGTGTGGCCAggcagttattagc
.c_hu	1540	GAACTATATCTGTAAGGCTCAAATAGTGTGGTCatacactgattact
rat	8051	gtgtCAGTGATGGTTGTGGCCTACCTCCCCGAATCATAAactacctACC
.c_hu	1590	gttcCAGTAATGATTGTAGCCTACCTCCCCGAATTACAAatogtccACC
rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
.c_hu	1640	CAGGAGTATTGTCAACAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
rat	8151	CCTTGGAATCCCCAAGCCaaAAGTCAOCTGGGAGACGCCAAGCACTCC
.c_hu	1690	CCTTGGGAGTTCCCCAAGCCagAAATCACATGGGAGATGCCGTGACCACTCC
rat	8201	CTGCTCTCAaaagcaacagcaagaaaacccCATAGAAGTGAGATGCTTCA
.c_hu	1740	CTTCTCTCAAcggcaagtaagagaggacaCATGGAAGTGAGCAGCTTCA
rat	8251	CCCACAAGGTACGCTgGTCATTTCAGAACTCTCCAAACCTCGGATTCCGGag
.c_hu	1790	CTTACAAGGTACCTTaGTCATTTCAGAAATCCCCAAACCTCCGATTCTGGga
rat	8301	tctATAAGTGCAGAGCTCAGAACCTACTTGGgacTGATTACGCAACAACT
.c_hu	1840	taTACAAATGCACAGCAAGAACCCTTGGtagTGATTATGCAGCAACG
rat	8351	TACATCCAGGTACTCTGACAGGAaggggagactaaaattoaacagaagt
.c_hu	1890	TATATTCAAGTAATCTGACATGAAataataaagtcacaacatctgggca

cDNA_rat	8401	ccacatccacagggTTTATTTTGGAGAAGTTAATCAAAGGCAGCCA
genomic_hu	1940	gaa-----TTTATTTTGGAGAAGTTAATCAAAGGCAGCCA
cDNA_rat	8451	TAGGCATGTAAATGagtcTGAATACATTACAGTATTAAATTTACAATGG
genomic_hu	1979	TAGGCATGTAAATGAattTGAATACATTACAGTATTAAATTTACAATGA
cDNA_rat	8501	ACATGCgatga-----GACTTGTAATGAAAGCATTGTGAAGTGAaaccg
genomic_hu	2029	ACATGCaaaataaaaagGACTTGTAATAAATGCATTATGAAGTGAatgata
cDNA_rat	8546	agtctctg-----TGGATCTCAAAGCAAACCTCTTAACTTAAGGCACTTTg
genomic_hu	2079	ctgatttatttaaTGGATCTCAAACAAACCTTTTAACTTAAGGCACTTTt
cDNA_rat	8591	ATTTTGCCAACAAATAATAACAAACattaagagaaaaaatgatacCACTA
genomic_hu	2129	ATTTTGCCAACAAATAACAATAACAaacattgaaacggtt----CACTA
cDNA_rat	8641	CGAAATAACAAACGGCTAATGCACCTGAATTctcagtaaaaaagacctttc
genomic_hu	2175	TAAATAACAAATGGCTAATGTACCTGAATTtttcagtaaaaaaatgaac
cDNA_rat	8691	tctcgctaacagttgCCAGCTGCCTCGTGTCTGTTTCTACCAATGTCAC
genomic_hu	2225	ttctaata-----CCAGTTGCCTAGTGTCCACCTCCTATCAATGTTAC
cDNA_rat	8741	AAACatogcacacagggTGAATGGAGTCAACGGGAAAGATTAAAGTTTGCG
genomic_hu	2268	AAGCATggcactca----GAACAGAGACAATGGAAATATTAAATCTGCA
cDNA_rat	8791	GTCTgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgatgtaaatattaccatcctgATGTATAAATATTTTG----TG

exon1 (2342-2397)

cDNA_rat	8791	GTCTgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgatgtaaatattaccatcctgATGTATAAATATTTTG----TG
cDNA_rat	8829	GTTTATAAACATTTTGATAAAACCGAAAAAAAAAAAAAAAAAAAAA
genomic_hu	2360	GTTTATAAATTTTTTGCTAAACCTACAGAAATAAGcactgaactgtc

(Genomic human OCP: SEQ ID NO:4)
(cDNA rat: SEQ ID NO:5)

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~231 7	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

Fig 12.

608 Human translated nucleotide sequence (ORF)

ATGAAGGTA AAAAGGCAGAGGAATCACCTGCTTGGTCTCCTTTGCTGT
GATCTGCCTGGTGCACCCCTGGGGGCAAGGCCTGTCTCGCCGCTGTG
CCTGTTATATGCCTACGGAGGTACACTGCACATTTGGGTACCTGACTTCCA
TCCCAGACAGCATCCCGCCAATGTGGAACGCATCAATTTAGGATACAAC
AGCTTGGTTAGATTGATGGAAAACAGATTTTTCTGGCCTGACCAAACCTGGA
GTTACTCATGCTTCACAGCAATGGCATTACACAATCCCTGACAAGACCTT
CTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAATAAAGTCC
GAAAACCTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTG
CACATGGACCACAACAATATTGAGTTTATAAACCCAGAGGTTTTTATGG
GCTCAACTTTCTCCGCCTGGTGCACCTGGAAGGAAATCAGCTCACTAAGCT
CCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAAATCTC
TTTCATTAAGTTCCTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAA
GAGATGGTCTCCTATATGCCTGACCTAGACAGCCTTTACCTGCATGGAAA
CCCATGGACCTGTGATTGCCATTTAAAGTGGTTGTCTGACTGGATACAGCC
AGATGTAATAAAAATGCAAAAAGATAGAAGTCCCTCTAGTGCTCAGCAGT
GTCCACTTTGCATGAACCTAGGACTTCTAAAGGCAAGCCGTTAGCTATG
GTCTCAGCTGCAGCTTCCAGTGTGCCAAGCCAACCATGACTCATCCCTG
AAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTTCTCTCT
CCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAG
GACATCACCCATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATA
CTTCATTTTCAACATTTTGGTGTGCAACATAGATTACGGTCACATTACGC
CAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCTCTGATACATAGAAA
GGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAG
GTGGCTCCTAAGCCTGAAGACATTTTACCAACATAGAGGCAGATCTCAG
AGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAGCTGAACA
GAACTGCCACCACATTCACTACATTACAGATCCAGTACTCCAGTGATGCT
CAAACTCACTTTACCAAGAGCAGAGATGAGGCCAAGTGAAACACAAATGGA
CTATGATTTCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTA
GGTGGAAACCGTTGGCCTGAACTGCCAGGCCAAGGAGACCCCAACCCACA
CGTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCA
GTGAGGATGGACGGATCCTAATAGACAAAAGTGGAATAATTGGAACCTCA
GATGGCTGATAGTTTGTACACAGGCGTATATCACTGTATAAGCAGCAATT
ATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGG
TCGAAGCCTATCAGGAAAATGGGATTCTACACACAGTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGC
TGGGTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAA
AGTTCTAAACAATGGCACATTAAGAATATTACAGGTCAACCCGAAAGACC
AAGGTTATTATCGCTGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGA
TTTTCCAAGTTTCAGTCAAGATGAAAGGACAAAGGCCCTTGGAGCATGAT
GGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCT
TAAGGAGCCACCAGGTGCACAACCTCCGACATCTGCTCTGATGGAGGCTG
AGGTTGGAAAACACACCTCAAGCACAAAGTAAGAGGCACAACCTATCGGGA
ATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGA
ATAGGAGGCATTTCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGG
GCGGCACTGTTGGAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAG
AAAATACCACAGTGAGCCACCCCAAGTGGTCAACCAACTCCCAACATA
CCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATT

TATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTG
CTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATAAAATTATGGC
ACAGAAATTCTCTCTGTTGTGAATTCACAAATACTACCACCTGAAGAACCC
ACAGATTTCAAACTGTCTACTGCTATTAATACTACAGCCATGTCAAAGAA
TATAAACCCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATT
CATCCACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTTCAGGACTCTG
ACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCAATAAC
AGTAAGGACTATGATCAAAGATGTCAATGTCAAATGCTTAGTAGCACCA
CCAACAAACTATTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACA
TCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCACTTCTATTCTCACAC
TACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGC
TCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTT
CAGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGCGGATTATC
AGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTTCAG
GTCAACAACCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCA
CAGTGCTCAATGTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCA
CTGCCACAGCAGCATTGTCTTTTCCAAGTGCTGCTCCCATCACCTTCCCCA
AAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCTTAGTCCAG
AATCCACTATTACTACTTGAGAACAACCCAGTGTAAGAGAAAACAACACC
CACAATAAAATATTTTCAGGACTGAAATTTCCCAAGTGACTCCAACCTGGT
CAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAATCACA
GTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGA
TTCAGTGATTACATCGTCACTTTTCAGGTGCTATCACCAGCCACCAATGAC
TATTATAGCCATTACAAGGTTTTCAGAAGGAAAATTCCTGGCAACAGA
ACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGAATCAACATAAA
GTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCTGCT
TTACCACAGAGACAAAGTTCCCTTTCCATTTACCCACACTTTCAACAAGT
GTGATGCAAAATTCATCTAATACTTGACTACCGCTCACCACACTACGACC
AAAACACACAATCTGGAAGTCTTCCAACAAAAGAGGAGCTTCCCTTCCC
ACCCCTTAACCCCTATGCTTCTTAGTATTATAAGCAAAAGACTCAAGTACA
AAAGCATCATATCAACGCAACAGCAATACCAGCAACAACCTCTACCTTC
CCTGCATCTGTCTACTTATGAAACCCAAACAGAGAGATCTAGAGACA
AACAATACAAAGAGAACAGGAGCCTCAAAAAGAAGAACAGGACTGACCCA
AACATCTCTCCAGACCAGAGTTCTGGCTTCACTACCCCACTGCTATGACA
CCTCCTGCTCTGGCATTCACTCATTCCCCACCAGAAAACACAACCTGGGATT
TCAAGCACAATCAGTTTTTCATTCAAGAAGCTCTTAATCTGACAGATGTGATT
GAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTC
TGAAACAACCTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAG
CATCATTAGACACTCCCATACCACCATCTTGAGCAGCAGTGCTACTCTAA
TGCCAGTTCCCATCTCCCTCCCTTTACTCAGAGAGCAGTTACTGACACAC
GTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCAC
GAATCCTCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACT
CACTTCATCTACCTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAAC
AACAGTTAAATCACAGAATTCCAAATTAACCTCATCTCCCTGGGCAGAAAT
ACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCAAAAAG
CCAGAAGTAAGCATGTTGGCTACTACAGGCTGTCCGAGGCCACCACTCT
TGTTTCAGATTGGGATGGACAGAAGAACAACAAAGAGTGACTTTGATA
AGAAACCAGTTCAAGAAGCAACAACCTTCCAAACTCCTTCCCTTTGACTCTT
TGTCTAGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCA
AGTTTTACTATTCCAGCTAACTCAGATGCCTTTCTTCCCTGTGAAGCTGTT

GGAAATCCCCTGCCACCATTTCATTGGACCAGAGTTCAGGACTTGATTTA
TCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCAATGGTACCCT
GTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTCCG
CATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGTCTGTGGTT
CCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCAT
TCCGGAAGCACTGTGGAAGTGCAGAGCAGAAGGTAGGCCAAGCC
CTACAGTTACCTGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCC
AGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGAACATTGGTCCTCCAC
AATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCA
GGTGGCCAGGATTCAGTCTGGTTAAAATACAAGTCATTGCAGCACCACC
TGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGCACCTGGGGTGAAA
GTTTAAAACTGCCCTGTACTGCAAAGGAACCTCTCAGCCCAGCGTTTACT
GGGTCTCTCTGATGGCACTGAAAGTGAACCATTAAGTTTACCAATTCCA
AGTTGTTCTTATTTCAAATGGGACTTTGTATATAAGAAAACCTAGCCTCTT
CAGACAGGGGCACTTATGAATGCATTGCTACCAGTTCCACTGGTTCGGAG
CGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGA
TAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTGGGGACAAATTA
CTACTGAACTGCTCAGCCACTGGGGAGCCCCAAACCCAAATAATGTGGAG
GTTACCATCCAAGGCTGTGGTCGACCACTGGAGCTGGATCCACGTCTACC
CTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTG
TACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCA
TGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTA
GAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCC
GGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGAT
CAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATAT
ACCCTTTTCAACAATGGAACCTTATACTTCAACAAAGTTGGGGTAGCGGA
GGAAGGAGATTATACTTGTATGCCAGAACACCCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGATAAGGCAGAG
TAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCTTGACT
GTGAGGTCACTGGGGATCCCCAAACCAAAAATATTTGGTTGCTGCCCTCC
AATGACATGATTTCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAACCTGCTCGATTCTGGAGAGTACGTATG
TGTAGCCCCGAAATCCCAAGTGGGGATGACACCAAAATGTACAACTGGATG
TGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACAAACAGAACTGTTA
TTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCT
GAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTT
CCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAAATGGAA
CCTTGGAATAGGAATGTGAGGCTTTCAGATTACAGCCGACTTATCTGTG
TGGCCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTA
CTGGAATGCTGAGAAGACCGACATTTAGAAATCCATTAATGAAAAAAT
AGTTGCCCAGCTGGGAAAGTCCACAGCATTGAATTGCTCTGTTGATGGTA
ACCCACCACCTGAAATAATCTGGATTTTACCAATGGCACACGATTTTCCA
ATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTTTTATCA
TTTCTAAACAACCTCGGGAGGATGCAGGAAAATATCGTGTGCAGCTAGG
AATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAA
GCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAGGCATCAGTGGAG
AATCTCTATCACTGCATTGTGTCTGATGGAATCCCTAAGCCAAATATCA
AATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAATTAATGGG
AAATACATATTGCATGACAATGGCACCTTAGTCATTAAGGAAGCAACAGC
TTATGACAGAGGAACTATATCTGTAAGGCTCAAAATAGTGTGGTCATA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

```

rat_cDNA      CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAGGAAA
human_5+3_corrected
mus_cDNA_5    -----CAGAAGGGTCCAGGAAA-GGAAA

rat_cDNA      GTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG
human_5+3_corrected
mus_cDNA_5    GTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGACATCGCTTCAGTGACTGAA

rat_cDNA      GCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCCTTCATAGCACTGAT
human_5+3_corrected
mus_cDNA_5    GCCAGGCAAAAGGAGCGGGGAAGGATTATATGTAGCCTGGGACGCTTCATAAACACTGAT

rat_cDNA      GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCCCTGGGCCTCGGAAAG---GA
human_5+3_corrected
mus_cDNA_5    GACGTGTTTGTGCAAAAGCAGCRAATTTGAGGAGAAACGCCCTGGGACGTCGGAAAGAAGGA

rat_cDNA      GTGATTGATTAGTACTTGCAGTTTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA
human_5+3_corrected
mus_cDNA_5    GTGATCGATTAGTACTTGTAGTTTAGGTGAGTTT--GAGAACTAACTAACCTATACTA

rat_cDNA      TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGCAGCAGGAAGCTTTGGTTAAT
human_5+3_corrected
mus_cDNA_5    TTGAGGGAGAGGAAGAGCATT-----CCAGCAGCAGCAGGAAGCTTTGGTTAGT

rat_cDNA      TTGGAATGATGATAGCATTAAATAACAGAAGCGCCTCCAGGTCTCTGAAGCTTCAGT
human_5+3_corrected
mus_cDNA_5    TTGGAATGTATGATACCATTAAATAACAGAAGCGCCTCCAGTTCTCTGAAGAGTCAGT

rat_cDNA      CCCCAGCTGAAAGCCAGAAAAGACTAAGCCCACTAAGCCTTTTGATCCCTTTGGAAGCA
human_5+3_corrected
mus_cDNA_5    CCCCAGCT-----A-GTGTAGCCTACTAAGCCTTTTGCTCCCGTTGGAAGCA

rat_cDNA      AAGAATTTTCCCTCCCTGGGGTGAAGACTCTCCTCAGAAGATTTCTGTCTCTGCCTATG
human_5+3_corrected
mus_cDNA_5    AAGAAGCTTCCCTCAATCAGGTGAAGGCTCTCCTCAGAAGATTTCTGTCTCTGCCTATG

rat_cDNA      TTACAAGAGGAATCAAAACCAAGACAGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAG
human_5+3_corrected
mus_cDNA_5    TTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGGATGCAGAAGAGAGGCAGGGAAG
*** ** * ***** *

rat_cDNA      TCAGCGGCTTGTGTCTCCCTCACTGCTGTCTGCTGGTGGTCAOCCCTGGGAGCAGGS
human_5+3_corrected
mus_cDNA_5    TCACCTGCTTGTGCTGCTCCTTTGCTGTGATCTGCCTGGTGGCCACCCCTGGGSCAAGS
TCAGCTGCTTGTGCTGATCTCCTCACTGCCATCTGCCTGGTGGTCAOCCCTGGGAGCAGGS
*** * ***** ** ***** * *** ***** * ***** *

rat_cDNA      CCTGTCTCGCGGCTGTGCTGCTATGTGCCCCAGAGGTGCACTGTACATTCGGTACC
human_5+3_corrected
mus_cDNA_5    CCTGTCTCGCGGCTGTGCTGCTATGTGCCCCAGAGGTACACTGCACATTTGGTACC
TCTGTCTCGCGGATGTGCTGCTATGTGCCCCAGAGGTGCACTGTACATTTGGGACC
***** ***** ** ***** * ***** *****

rat_cDNA      TGACCTCCATCCAGATGG-CATCCCGCCCAATGTGGAACGAATAAATTTAGGATATAAC
human_5+3_corrected
mus_cDNA_5    TGACTTCCATCCAGACAG-CATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAC
TGACCTCCATCCAGACGGGCATCCAGCCCAATGTGGAACGAGTCATTTAGGGTATAAC
**** ***** * ***** ***** * ***** **

```

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 AGCCTTACTAGATTGACAGAAAACGACTTTGATGGCCTGAGCAAACTGGAGTTACTCATG
 AGCTTGGTTAGATTGATGGAACAGATTTTCTGGCCTGACCAAACTGGAGTTACTCATG
 AGCCTCACTAGATTGACAGAAAATGACTTTCTGGCCTGAGCAGACTGGAGTTACTCATG
 *** * ***** ** * ***** *

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 CTGCACAGTAATGGCATTACACAGATCAGTGACAAGACCTTCTCGGGCTTGCACTCTTG
 CTTCACAGCAATGGCATTACACAAATCCCTGACAAGACCTTCTCAGATTGTCAGGCTTG
 CTGCACAGCAATGGCATTACACAGATCAGTGACAAGACCTTCTCGGGCTTGCACTCTTG
 ** ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 CAGGTCCTAAAAATGAGCTATAACAAAGTCCAAATCATTGGGAAGGATACTTTCTACGGA
 CAGGTCCTAAAAATGAGCTATAAATAAGTCCGAAAACCTCAGAAAGATACTTTTATGGC
 CAGGTCCTAAAAATGAGCTATAACAAAGTCCAAATCATTGAGAAGGATACTTTGTATGGA
 ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 CTCGGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCCTGAG
 CTCAGGAGCTTGACACGATTGCACTGGACCAACAATATTGAGTTTATAAACCAGAG
 CTCAGGAGCTTGACCCGGTTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAG
 *** ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 GCGTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACAAG
 GTTMTTATGGGCTCAACTTTCTCCGCTGGTGCACTTGGGAAGGAAATCAGCTCACTAAG
 GCGTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACAAG
 * ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 CTCCATCCAGACACATTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTTTCATT
 CTCCACCCAGATACATTGTCTCTTTGAGCTACCTCCAGATATTTAAAATCTCTTTCATT
 CTCCATCCAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCTTTCATT
 ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 AAGTACCTGTTCTGTCTGATAACTTCCT-GACCTCCCTCCCAAAGAAATGGTCTCCTA
 AAGTTCCTATACTGTCTGATAACTTCCT-GACCTCCCTCCCAAAGAGATGGTCTCCTA
 AAGNACCTGTACTGTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTC
 *** ** * ***** *

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 CATGCCAAACCTAGAAAGCCTGTATTGCAATGGAAACCCATGGACCTGTGACTGCCATTT
 TATGCCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTT
 TATGCCAAACCTAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTT
 ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 AAAGTGGTTGTCTGAGTGGATGCAGGGAACCCAGATATAATAAATGCAAGAAAGACAG
 AAAGTGGTTGTCTGACTGGATACAGGNNNNOCAGATGTAATAAATGCAAAAAGATAG
 AAAGTGGTTGTCCGAGTGGATGCAGGGAACCCAGGTA-ACTATCTGT-TTGTGTG
 ***** **

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 AAGCTCTTCAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
 AAGTCCCTCTAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
 TTTCTTTTTTATARKAGTATTTCTCAATTTCAATTTAGAAATGATATCCCAAAGTC-
 * * * * *

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 GACCCCTTGCTATGTTACCATCTGGAGCTTTCTATGTACAAAGCCACCATTTGATCCAT
 AGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAGTGTGCAAGCCACCATTTGACTCAT
 -CCCATACCTCCCTCCCA-----CTTCCCTACCTACCCATTC-CCATTTTGGC
 ** * *

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 CACTGAAGTCAAGAGCCTGGTTACTCAGGAGCAATGGATCTGCTCCCTCCACTCACCTC
 CCTGAAATCAAGAGCCTGACTATTCTGGAAGACAGTACTCTGCTTTTCATCTCTCCCC
 CCTGGCATTCCCC-----
 * * *

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 AAGATTTATAGAACCTTTGGCTCCTTGTCTTTGAACATGACANANNTNTCTGGAAATA
 AAGGTTTCATGGCACTTTGGCTCCCTCATTGAAATGACAGATCAGTCTGGAAATG

rat_cDNA
 human_5+3_corrected
 mus_cDNA_5
 AGGCCGACATGGTCTGTAGTATCCAAAAGCCATCAAGGACATCACCAACTGCATTCACTG
 AAGCTAACATGGTCTGCAATATCCAAAAGCCCTCAAGGACATCACCCATTGCATTCACTG

rat_cDNA human_5+3_corrected mus_cDNA_5	AAGAAAATGACTACATCATGCTAAATGCGTCATTTTCCACAAATCTTGTGTGCAGTGTAG AAGAAAATGACTACATCGTGTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ATTATAATCACATCCAGCCAGTGTGGCACTTCTGGCTTTATACAGTGACTCTOCTCTGA ATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTOCTCTGA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGG TACTAGAAAGGAGCCACTTGCTTAGTGAACACCCGAGCTCTATTACAAATATAAACAGG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TGGCTCTTAGGCCCTGAAGACATTTTACCASCATAGAGGCTGATGTCAGAGCAGACCCCTT TGGCTCCTAAGCCTGAAGACATTTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TTTGGTTCCACACAGAAAAATTTGCTTGCAGCTGAACAGAACTGCCACCACACTTAGCA CTTGGTTAATGCAAGACCAAATTTCCCTTGCAGCTGAACAGAACTGCCACCACACTTAGTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CATTACAGATCCAGTTTTCCTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAG CATTACAGATCCAGTACTCCAGTGAATGCTCAAATCACTTTACCAAGAGCAGAGATGAGGC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CGGAGAGACTCAAATGGACCATGATCCTGATGATGAACATCCCAAAGTGGAAACGCACTG CAGTGAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACATACTG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TCTTGGTTGGCGGCACTATTGCCCCGAGCTGTCCAGGCAAGGGGACCCCTTCACTCACT TCTTGGTAGGTGGAACCGTTGGCCTGAAGTCCCAGGCCAAGGAGACCCCAACCCACAGC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTACGTTAGCGAGGATGGGC TGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCAGTGAGGATGGAC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGATGCAG GGATCCTAATAGACAAAGTGAAGTGGAACTCCAGATGGCTGATAGTTTTCACACAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATACAGGATAA GCGTATATCACTGTATAGCAGCAATATGATGATGCAGATATCTCACCTATAGGATAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA CTGTGGTAGAACCTTTGGTCCAGGCTATCAGGAAAAATGGGATTATCACACAGTTTTCA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CGGGTGAGAGCGCTGACCTTCCATGCTTTCCAGGGTGTTCAGATGCTTCTATTAGCT TTGGTGAAACACTTGATCTTCCATGCCATTTCTACTGGTATCCCAGATGCTTCTATTAGCT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GGATTCTTCCAGGGAACTGTGTTCTCTCGCCATCAAGAGACAGGCAAATCTTAACA GGGTTATTCCAGGAAACATGTGCTCTATCAGTCATCAAGAGACAGAAAGTTCTAACA -----

rat_cDNA human_5+3_corrected mus_cDNA_5	ATGGGACCTTAAGAATATTACAGGTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG ATGGCACATTAAGAATATTACAGGTACCCCCGAAAGACCAAGGTTATTTATCGCTGTGTGG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTGCCAACCCATCAGGGGCGGACTTTTCCAGTTTAAAGTTTCAGTTCAAAAGAAAGGCC CAGCCAACCCATCAGGGGTGATTTTGTATTTTCCAAGTTTCAGTCAAGATGAAGGAC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AAAGGATGGTTGACATGACAGGGAGGCAGGTGGATCTGGACTTGGAGAACCCAACTCCA AAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTGATGAGTCCAATCCTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGTTTCCCTTAAGCAGCCAGCATCTTTGAAACTCTCTGCATCAGCTTTGACAGGGTCAG TTGCTCATCTTAAGGAGCCACAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AGGCTGGAAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC AGGTTGGAAAACACACCTCAAGCACAACTAAGAGGCACAACTATCGGGAATTAACTCTCC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GGCGGCGTGGGGATTCCAGCTCCGGCGATTGAGGAGCATAGGAGGCAGCTCCCTCTCT AGCGACGTGGAGATTCAACACATCGAGCTTTAGGGAGAAATAGGAGGCATTCCCTCCCT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTGCTCGGAGAAATGACCCGCAACCTGGGCGAGCACTTCTAGAAAAGCCAAAAGAAAT CTGCTAGGAGAAATGACCCACACATTTGGGCGGCACTGTTGGAGAAAGCTAAAAGAAATG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTGTGCCAAAAAGCAAGAAAATACCAAGTAAAGCCAGTGGCACTGGCTGTTCCCTCTG CTATGCCAGACAAGCGAGAAAATACCAAGTGAAGCCACCCCAAGTGGTCAACCAACTCC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TGGAACCTCACTGACGAGGAAAAGGATGCTCTGGCATGATTCCTCCAGATGAAGAAATCA CAACATACCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATTTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TGGTTCTGAAAACCTAAGGCTTCTGGTGTCCAGGAAGGTCAACAACTGCTGACTCTGGAC TGGTCCCGCCCACTAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CAGTAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC CAATATCTGATAGTCCATGACAAACATAAATTTATGGCACAGAAGTCTCCGTTGTGAAT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CACAAACACTACAACTGAGCACCTTCCTGATTTCAAAATATTTAGTGTAAACAAAGGTA CACAAATACCTACCACTGAAGAACCCACAGATTTCAAACCTGTCTACTGTATTAATACTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CAGCTGTGACAAAGAGTATGAACCATCCATAGCAAGCAAAATAGAAGATCAACCAACC CAGCATGTCAAGAAATATAAACCACCAATGTCAAGCCAAATAGAGGCAACCAACATC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AAAACCCAACTCATTATCTTTCCATC-----AGTAGCTGAAATTGAGATTTCTG-CT AACATTCACTCACTGTCTTTCCACTGCTAATTGGAGCAACTGAATTTCAGGACTCTGACA -----

rat_cDNA human_5+3_corrected mus_cDNA_5	CAGGCA---GGAAGAGCATCTTCCAAAGTGCAACCCCTGTAACAGGGGAAACATGGCT GAGGGAAGAGGAAGAGAGCATTTCC--AGTA-ACCCCCAATAACAGTAAGGACTATGATC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ACCTATGGCCATACCAACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTGC AAAGATGNTCAATGTCAAANATGCTTAGTAGCACCACCAACAAA-CTATTA-----TTAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AGCCATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCAGCAGAC AGTCAGTAAATACCACAAATAGTCAT-----CAGACATCTGTAGAGAAAGTGAGTGAAC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCCCTAGCTTCTCCAGTCACCCTT CCAGGCACAATCACTTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CAGGTTCAACACACCCTGCGCTCTTTATTTACATTCCTAGAAACAACAATACAGGTA CAGATCCACACACAGCTGCTCATTTCTCAGTTTCCGATCCCTAGANNNAATAGTACAGTTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ACTTCCCTTGTGCCAGGCACTTGGGAAGAGAGAGGACAATTGGAGCAGAGGGGAGAGTTA ACATCCCGCTGTTCAGACGCTTTGGGAGGCAGAGGAAAATTGGCGAAGGGGGGGGATT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AAAACCCATAGAAACCCAGTTCTCCGACGGCATAGACACAGGACTGTGAGGCCAGCAA TCAGCCCATATAGAACTCCAGTTCTGCGAOGGCATAGATACAGCATTTTCAGGTCAACAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TCAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCAGCCACAGAGTACCCCTGGGATGT CCAGAGGTTCTTCTGAAAAAAGCACTACTGCAATTCTCAGCCACAGTGCTCAATGTGACAT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GCCACACATGTCTTCCGAGAGGGGCTCAGTGGCTACTGCAGCACTGTCACTTCCAA GTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTCCAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GTTCAATCCCACTGCGCTCCCAAACTAATAATGTGGGGTCATAGCAGAAGAGTCTA GTGCTGCTCCCATCACTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CCACGTGTGTCAAGAAACCACTGTACTATTTAAGGACAAACAAAATGTAGATATTGAGA CAACTCTAGTCCAGATCCACTATTAATACTTGTAGAACAAACCCAGTGTAGANNNGAAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TAATAACAACCACTACAAAATATTCGGGAGGGGAAAGTAACCAAGTGATTCTTACGGAAG NNACACACCCACAATAAATAATTCAGGACTNGAAATTTCCCAAGTGACTCCAACTGGTG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CAAGCATGACTTCTGTCCAACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTG CAGTCATGACATATGCTCCACATCCATACCCATGGAAAAAATCACAAGTAACAGCCA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GTCACCTGAGCATGCTGGGAACATCCAAACTGGGAAAGATTCAAGTGAAACACACAC GTTACCCACGTGTGTCTAGCAACCAATGAAGCTAAAAGAGATTCAAGTATTACATGCTCAG -----

TTTCCAGCCCCCTCAGCACACCTTCAATAACAA-----CAAGCACAAATTCTCAAGA
TTTCAGGTGCTATCACCAAGCCCAATGACTATTATAGCCATTACAAGGTTTTCAAGAA

GGAAAACTCCCTTGCACCAGATCTTTGTAAATAACCAAGAAGGAGGGGATGTTAAAGA
GGAAATTCCTGGCAACAGAACTTTGTAAATAACCAATAACCAAAAGGCAGATTAAAGGA

ATCCATATCAATTGGTTTACAAAAGAACCCAGCCGCAAGGTTCCCAAATAGCTCCTC
ATCAACATAAGTTAGTTTACAAAAGACACAGCTGTGATGCTTCCTAAACATCTCCTG

TTTACCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGACA
CTTTACC-ACAGAGACAAAGTTCCCTTTCCATTTCACCCACACTTTCAACAAGTGTGATG

CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAGGGCACTGAAGTAGTATCAGGT
CAAATTCATCTAATACTTGACTACCGCTCACCACACTACGACCAA--ACACACAA-T

GCCAGAAGTCTCTCAGCAGGGAGAGCAGCCCTTCAOCAAATCTCTCCAG-TGCTTCC
CCTGGAAGTCTTCCACAAAGAGGAGCTTCCCTTC-CCACCCCTTAACTATGCTTC

TAGCACATAAGCAAGAGATCTAATACATTAAAGTTCTTGTCAACGGAAACCCCCACAGT
TAGTATTATAAGCAAAGACTCAAGTACAAAAGCATCATATCAACGCAACAGCAACCGC

GACAAGTCCTACTGCTACTGCATCTGTATTATGTCTGAAACCCACGAACAGATCCAA
AACAACTCCTACCTTCCCTGCATCTGTCTACTTATGAAACCCAAACAGAGAGATCTAG

AGAGCTAAAGACCAATAAAGGGGCTCG--GAGGAACGAAACACGCAAAACACCAC
AGCACAAACAATACAAAGAGAAGGACCTCAAAGAGAAGACAGGACTGACCCAAACATCTC

CCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACACAGCTGATAOCCOCTT
TCCAGACCAGAGTTCTTGGCTTCTACTACACCCACTGCTATGACNAGCTOCTNNGCTCTNNN

GGCTTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAGTGCAAGTTGCTTATCA
NGCATTCACTCATTCCCCACCAGAAAACACAACTGGGATTTCAAGCACAAATCAGTTTTC

CTCAACAACCTCTCTCTGGCCA---TAACTGAACTGTTTGAG-AGGTAC--ACCCAGAC
TTCAAGAAGCTCTTAATCTGACGATGTGATTGAGAACTAGCCCCAAGCAAGTAAGTCTAGAC

TTTGGGAATACAAACAGCTTTGGAAACAAAGTTGTTGAGCAAATCACAGGAGAGTACCAC
TTTGAAGAGCACAAATTGCTTCTGAAACAACTTTGTCCAGCAAATCACACCAGAGTACCAC

ASTGAAAAGAGCCTCA--GACAC--ACCACCACCCTCCTCAGCAGTGGGGCGCCCC
 AACTAGGAAAGCATCATTAGACACTCAACCAACCACCATTCCTTGAGCAGCAGTGCTACTCT

rat_cDNA human_5+3_corrected mus_cDNA_5	AGTGCCTCACTCCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTTACAGACAGCAAAGTCAC AATGCCAGTTCCCATCTCCCTCCCTTTACTCAGAGAGCAGTTACTGACAAAGTGGCGAC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ATCAGCTTTCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA TCCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCAGGAATCCTCAAGGCACAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TACAGATCTGCAGCAACCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATAACTGGAAC TCCNNNNNNNCA--AATGCCAAGTTCA-----CNAATTGNGAACNNNNNACTCNNNN -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CACTGACTCTCCCTCTAACTCTGTTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA NACTTCATCTACNTCTARTCTGTTACATTCTACTCCCATGCCAGCACTAACAACAGTTAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ACCACAGAATTCTAAATGGAAGCCCTCTCCTGGCCAGAACACAAATATCAGCTCAAGTC ATCACAGAATTCCAAATTAACCTCATCTCCTGGGCAGAAATACCAATTTTGGCACAACCC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ATACTCCGAAACCTTTGAGAAGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCCTCAG ATACTCAGACATTGCTGAAAAGGCCAAAAGCCAGAAGTAAGCATGTTGGCT-ACTACAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	-CCTTCAGAGGCCAGCACTCATGCCCTCACACTGGAATACACAGAAGCATGCAGAAAAGA GCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAGAACACAAAGAGA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGTTTTTGATAAGAACTGGTCAA-AAAC--CRACTTCCAAACATCTGCTTACGTCT GTGACTTTGATAAGAAACCAAGTTCAAGAAGCAACACTTCCAAACTCCTTCCCTTTGACT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTCTACCTTACACTCTATTGAAAAGGCCAAGATAATTGGAGGAAAGGCTGCAGCTTTA CTTTGTCTAGGTATATATTGAAAAGGCCAGATAGTTGGAGGAAAAGCTGCAAGTTTAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CAGTCCAGCTAATTCAGACGTTTTTCTTCTGTTGAGGCTGTTGGAGACCCACTGCCCA CTATTCAGCTAATCAGATGCCCTTCTTCCCTGTGAAGCTGTTGGAAATCCCTTGCCCA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TCATCCACTGGACCAGAGTTTCATCAGGANTTGAATATCCCAAGGGACACAGAAAAGCC CCATTCAATTGGACCAGAGTNNNNTCAGGACTTGATTATCTAAGAGGAAACAGAAATAGCA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GGTTCACGTTGCTTCCCAATGGCACCTTGTCATCCAGAGGGTCAGTATTCAGGAACCTG GGGTCCAGGTTCTCCCCAATGGTACCTGTCCATCCAGAGGGTGGAAATTCAGGAACCGG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCAATTTTATGTCCTTTGT GACAGTACTTGTGTTCCGATCCAATCTGTTGGCAGACACACCTTCATGTCACCTTGT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTGTGGTTTTTTTACCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCAGTTTCACT CTGTGGTTTCTATCTCTCCAGGATCTGGAGAGAGTACCAAGAGATCAGAGTTCAAT -----

rat_cDNA human 5+3 corrected mus_cDNA_5	TTGGAAGTACTGTGGAACTAAAGTGCAGAGTGGAGGSTATGCCAGGCCTACGGTTTCCT CCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCT
rat_cDNA human 5+3 corrected mus_cDNA_5	GGATACTTGCAAACCAACGGTGGTCTCAGAAACGGCCAAAGGGAAGCAGAAAGGTCTGGG GGATTCTTGCAAAACCAACAGTTGTCTCAGAATCATCCAGGGGAAGTAGGCAGGCTGTGG
rat_cDNA human 5+3 corrected mus_cDNA_5	TACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA TGACGGTTGACGGAACATTGGTCTCCACAACTCTAGTATTTATGACCGTGGCTTTTACA
rat_cDNA human 5+3 corrected mus_cDNA_5	AGTGTGTGGCCAGCAACCCATCTGSCCAGGATTCACTGTTGGTTAAGATACAAGTCATCA AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAATACAGTCATTG
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGCTCCCCCTGTCAATTATAGAGCAAAAGAGGCAAGCCATCGTTGGGGTTTTAGGTGGAA CAGCACCACTGTTATTTCTAGAGCAAGGAGGCAAGTCATTGTAGGCACCTTGGGGTGAAA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTTTGAAACTGCCCTGCACCTGCAAAAGGAACCTCCCGAGCCTAGTGTTCACCTGGGTCTTT GTTTAAAACTGCCCTGTACTGCAAAAGGAACCTCCTCAGCCAGCGTTTACTGGGTCTCTCT
rat_cDNA human 5+3 corrected mus_cDNA_5	ATGATGGGACTGAACTAAAAACATTGCAGTTGACTCATTCAGATTTTTCTTGATCCAA CTGATGGCACTGAAGTGAACCAATTACAGTTTACCAATTCCAAGTTGTTCTTATTTTCAA
rat_cDNA human 5+3 corrected mus_cDNA_5	ATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGTGGGGGCACTTATGAGTGCATTG ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATTG
rat_cDNA human 5+3 corrected mus_cDNA_5	CCACAGCTCCTCAGGCTCAGAGAGAAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA CTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAATCCCCAGGATAGAACTGCCCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT CCAGCCCCAGGATAGAACTGCATCCAGAAAAGGACTGAAGTGAATTTGGGGACAAAT
rat_cDNA human 5+3 corrected mus_cDNA_5	TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGCTGCCAT TACTACTGAACTGCTCAGCCACTGGGGAGGCCAAACCCCAATAATGTGGAGGTTACCAT
rat_cDNA human 5+3 corrected mus_cDNA_5	CCAGGCTGTCTATCGACCACTGGCAAGAAATGGGCAAGCCGAATCCAGTCTACCCAAATG CCAGGCTGTGTCTGACCACTGG-----GCAGCTGGATCCAGCTCTACCCAAATG
rat_cDNA human 5+3 corrected mus_cDNA_5	GATCCTTGGTGGTGGGTGAGTGAACGGAAGACGCTGGTGACTACTTATGTGTGGCAA GATCCCTGTTTATTGGATCAGTAACAGAAAAGACAGTGGTGTCTACTTGTGTGTGGCAA
rat_cDNA human 5+3 corrected mus_cDNA_5	GAAACAAAATGGGAGATGACCTAGTCTGTATGCATGTCCGCCTGAGATTGACACCTGCCA GAAACAAAATGGGGGATGATCTGATACGATGCATGTAGCCTAAGACTGAAACCTGCCA

rat_cDNA human_5+3_corrected mus_cDNA_5	AAATTGAACAGAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAACTTG AAATTGACCACAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAACTAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCTGATGGGACAG ATTGCAAGCTTCGGCTCCCAAGTCCAGAGATATCTGGAGTTTGCTGATGGAACTA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT TGATCAACATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TCCACAATGGAACTTGTATTTCAACAAGCTTGGGATGGCAGAGGAAGGAGATTATATCT TCAACAATGGAACTTTATACCTCAACAAGCTTGGGATAGCGAGGAAGGAGATTATACTT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GCTCTGCCAGAACACCTTAGGGAAGATGAGATGAATGTCCACCTAACAGTTCTAACAG GCTATGCCAGAACACCTTAGGGAAGATGAATGAAGGTCCACTTAACAGTTATAACAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CCATCCACGGATAAGGCAAGCTACAGACCCATGAGGCTCAGGGCTGGAGAACAG CTGCTCCCGGATAAGGCAGAGTAACAAACCAACAGAGAAATCAAGCTGGAGACACAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTGTCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTATTTTGGTTGCTGCCTT CTGTCTTGACTGTGAGGTCACTGGGATCCCAACCAAAATATTTTGGTTGCTGCCTT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CCAACAATGTCAATTCATTTCTCCAATGACAGGTTCAATTTCTGCCAATAGACCTTTGT CCAATGACATGATTTCTTTCTCCATGATAGGTACACATTTCTGCCAATGGGTCTTTGA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CCATCCATAAAGTGAAACCACTTGACTCTGGGACTATGTGTGCGTAGCTCAGAACTCTA CCATCAACAAGTGAAACTGCTCGATTTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGGGGATGACACTAAGACATACAACTGGAATTTGTCTCTAAACCTCCATTAACTCAATG GTGGGGATGACACCAAAATGTACAACTGGATGTGGTCTCTAAACCTCCATTAACTCAATG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GCCTGTATGCAACAAGACTGTTATTAAAGCCACAGCATTCGGCACTCCAAAAATACT GTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCAAAAAACAAT -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TTGACTGCAGAGCAGATGGGATCCCATCTTCCAGGTCACGTGGATTATGCCAGGCAATA TTGACTGCAGAGCTGAAGGGACCAATCTCTGAAATCATGTGGATCATGCCAGGCAATA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TTTTCCTGCCAGCCCCATACTATGGAAGCAGAAATCACAGTCCATAAAATGGAACTTTGG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AGATGAGGAACATCCGCTTTCTGACTCTGCGGACTTCACCTGTGTGGTTCCGAGCGAGG AAATTAGGAATGTGAGGCTTTCAATTCAGGCTTATCTGTGTGGCCCGAAATGAAG -----

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GAGGAGAGAGTGTGTTGGTAGTCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT
GTGGAGAGAGCGTGTGTTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAAC
TTAGAAATCCATTTAATGAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTCTGTGGATGGGAACCCGCCAAGTGAATTAAGTGGATCTTACCTGAAGGCACACAGT
GCTCTGTGATGGTAACCCACCAAGTGAATTAAGTGGATTTTACCAATGGCACACGAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTGCTAACAGACACACAATTCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT
TTTCCAATGGACCAAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTTTTATCATT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
CTAAACAACTCGGGAGGATGCAGGAAAAATATCGCTGTGCAGCTAGGAATAAGTTGGCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACATCGAGAACTCATCTGTAGAGATTGGGCAGAAGCCAGTCATTCTGACATAAGAAC
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGGGATGGTGRAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTCTGATGGGA
CAGGGACAGTAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTGATGGAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCTCAAG
TCCCTAAGCCAAATATCAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCTCAAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGGATGGAAATACATACTGCATGAAATGGCAGCTGGTCATCAAAGCAACACAGCTC
TTAATGGGAATACATATTGCATGACAAATGGCACCTTAGTCATTAAAGAACCAACAGCTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACGACCAAGGAAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCAGTTATTAGCG
ATGACAGAGGAACTATATCTGTAGGGCTCAAAATAGTGTGGTCATACACTGATTACTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGTCACTGATGGTTGTGGCTACCGTCCCCGAATCATAAACTACCTACCCAGGAACATGC
TTCCAGTAATGATTGTAGCTACCGTCCCCGAATTACAAATCGTCCACCCAGGAGTATTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCCAAGCCAA
TCACCAGGACAGGGGAGCGCTTTAGCTCCACTGTGTGGCCTTGGGAGTCCCCAAGCCAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGTCACCTGGGAGACGCAAGACACTCCCTGCTCTCAAAGCAACAGCAAGAAAAACCC
AAATCATATGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATAGAAGTGAGATGCTTCAACCAAGGTACGCTGGTCATTGAGATCTCCAAACCTCGG
ATGGAAGTGAGCAGCTTCACTACAGGTACCCTAGTCATTGAGATCCCCAAACCTCGG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATTCCGGAGTCTATAAGTGCAGAGCTCAGAACCTACTTGGGACTGATTACGCAACAACCTT
ATTCTGGGATATACAAATGCACAGCAAGAACCACCTTGGTAGTGATTATGCAGCAACGT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACATCCAGGTAAGTCTGACAGGAAGGGGAGACTTAAATTCACAGAGTCCACATCCACA
ATATTCAAGTAATCTGACATGAA----ATAATAAGT-CAACAA----CATCTGGGCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGGTTTATTTTGGAGAGAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA
GAATTTATTTTGGAGAGAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAATTTGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATACATTTACAGTATTAAATTTACAATGGACATGCGA--TGA---GACTTGTAAATGAAA
ATACATTTACAGTATTAAATTTACAATGAACATGCAAAATAAAAGGACTTGTAAATTAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCATTGTGAAGTGAA---ACCGAGTCTCTG--TGGATCTCAAAGCAAACCTCTTAACCTTAA
GCATTATGAAGTGATGACTGATTTATTTAATGGATCTCAAAACAAACTTTTAACCTTAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGCACTTTGATTTTGCCAACAAATATAACAAACATTAAGAGAAAAAATGATCCACTAC
GGCACTTTTATTTTGCCAACAAATAACAATAAACAA---ACATTGAACGGTTCACTAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAAATAACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTTCTCTGCTAAC
AAAATAACAAATGGCTAATGTACCTGAATTTTTCAGTAAAAAAA--TGAACCT-CTAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGTTGCCAGCTGCCTCGTGTCTGTTTCCACCAATGTCACAAACATCGCACACAGGGGTGA
A---CCAGTTGCCTAGTGTCCACCTCCTATCAATGTTACAAGCATGGCACTCAG---A

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGGAGTCAACGGGAAGATTAAAGTTTGCAGTCTGTGTAATCTCAATGTACAAATATTC
ACAGAGCAATGGAAAAATTAATCTGC-----AATCT--ATGTATAAATATTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGTCNCTGGTTTATAAATTTT-GATAAAACCGAAAAAAGAAAAAAGAAAAAAGAAAAA
TGT---GGTTTATAAATTTTGTCTAAACCTACAGAAAAATAG-----

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAA

(rat_cDNA: SEQ ID NO:7)

(human_5+3 corrected: SEQ ID NO:8)

(mus_cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVRGREVSGLLISLTAVCLVVPFSGRACPRRCACYVPTVEVHCTFRYLTSIPDGIPANVE
human_5+3_corrected	MRVKGGRGITCLLVSEFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPPNVE
mouse_5_corrected	MQKRGREVSGLLISLTAVCLVVPFSGRVCFPRRCACYVPTVEVHCTFRDLTSIPD-GPANVE
rat	RINLGYNSLTRLTENDFDGLSKLELIMLHSHNGIHRVSDKTFSGLSQVLKMSYNKVQII
human_5+3_corrected	RINLGYNSLVRIMETDFSGLTKELELIMLHSHNGIHTIPDKTFSDLOALQVLKMSYNKVRL
mouse_5_corrected	RVNLYNSLTRLTENDFDGLSKLELIMLHSHNGIHRVSDKTFSGLSQVLKMSYNKVQII
rat	RKDTFFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTEFVLSYLQ
human_5+3_corrected	QKDTFFYGLRSLTRLMDHNNIEFINPEVYGLNTRLVHLEGNQLTKLHPDTEFVLSYLQ
mouse_5_corrected	EKDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTEFVLSYLQ
rat	IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLYHGNPWTCDCHLKNLSEWMOGNPDI
human_5+3_corrected	IFKISFIKFLYLSNFLTSLPQEMVSYMPDLDSLYLEGNPWTCDCHLKNLSDWIQ--PDV
mouse_5_corrected	IFKTSFIKXYLYDNF-TSLPKEMVSSMPNLESYLYHGNPWTCDCHLKNLSEWMOGNP--
rat	IKCKKDRSSSSPQOCPLCMNPRIKGRPFAMVPSGAFLCTKPTIDPSLKSLSLVTQEDNG
human_5+3_corrected	IKCKKDRSPSSAQOCPLCMNPRTSKGKPLAMVSAAFQCAKPTIDS6LKSLSLTILEDSS
mouse_5_corrected	-----
rat	SASTSPQDFLEFPFGLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
human_5+3_corrected	SAFISPGGFMAFPFGLSLNMTDQSGNEANMVCSIQKPSRTSPTAFTEENDYIVLNTSFST
mouse_5_corrected	-----
rat	NLVCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHQIPVWQILALYSDSPLILERSHLLSETPOLYKYKQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADPFWFQEQKIVLQINRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWIMILMNN
human_5+3_corrected	DLRADPSWLMQDQISLQINRTATTLSTLQIQYSSDAQITLPRAMRPVKIKWTMISRDNN
mouse_5_corrected	-----
rat	PKLERTVLVGSTIALSCPGKGDPSPHLENLLADGSKVRAPYVSEDGRILIDKNGKLELQ
human_5+3_corrected	TKLEHTVLVGSTVGLNCPGQGDPTPHVWLLADGSKVRAPYVSEDGRILIDKSGKLELQ
mouse_5_corrected	-----
rat	ADSF DAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDPCLSTGV
human_5+3_corrected	ADSEDTGVYHCISSNYDDADILTYRITVVEPLVEAYOENGIHTVFIGETLDPCHSTGI
mouse_5_corrected	-----
rat	PDASISWILPGNTVFSQPSRDRQILNNGTLRILOVTPKQOGHYQCVANPBGADFSSTKV
human_5+3_corrected	PDASISWVIFGNVLYQSSRDKKVLNNGTLRILOVTPKQGYRYRCVANPBGVDLIFQV
mouse_5_corrected	-----
rat	SVQKKGQRMVEHDREAGGSGLGFPNS8VSLKQFASLKLASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVRMKGQRPLEHDGETEGSGGLDESNPITAHKEPPGAQLRTSALMEAEVGRHTSGSTGRHN
mouse_5_corrected	-----
rat	HRDLIHRRRGDSTLRFRFHRRQLPLSARRIDPQWAAALLEKAKKN5VPAKQENTTVKPV
human_5+3_corrected	YRELTLOQRGDSTHRRFRFENRRHFP5ARRIDPQWAAALLEKAKKNAMPDKRENTTVSPP
mouse_5_corrected	-----

rat	PLAVPLVELTDEEKDASGMIPDEEFMVLTKASGVPGRSPTADSGPVNHGFMSTIASGT
human_5+3_corrected	PVVTQLFNIPGEEEDSSGMLALHEEFMVPATKALNLPARTVTADSRITSDSPMTNINYGT
mouse_5_corrected	-----
rat	EVS-TVNPQTLQSEHLPDFKLFSVINGTAVTKSMNPSLASKIEDTTNQNPIIIFP---SV
human_5+3_corrected	EFSPVVNSQILPPEEPTDFKLSTAIKTTAMSKNINPTMSSQIQGTTNQHSSSTVFPILLGA
mouse_5_corrected	-----
rat	AEIRDSAQAGRAS--SQSARPVTCGMATYGHNTYSSFTSKASTVLQPINPTESYGPQI
human_5+3_corrected	TEFQSDOMGRGREHFQSRPPIVETMIKDVNVKMLSSTTNKL--LLESVNTTNSH--QT
mouse_5_corrected	-----
rat	PITGVSRPSSSDISSHTTADPSFSHPSGSHTTASSLEHIPPNNNTGNFPLSRHLGRERT
human_5+3_corrected	SVREVSEPRNHFYSHTTQILSTSTFPSPDHTAHSQFPIPRNS-TVNIPLFRFRGRQRK
mouse_5_corrected	-----
rat	IWSRGVKNPHRTFVLRHRHRTVPAIKGPANKNVSQVPATEYPMCHTCPSAEGTLVA
human_5+3_corrected	IGGRGRIISPVRTFVLRHRYSIFRSTTRGSSEKSTAFSATVNLNVTCLSLPRERLTTA
mouse_5_corrected	-----
rat	TAALSVPSSSHSALPKTNVGVIAEESTTVVKKPLLLFKDRQNVDIETTTTKYSGGES
human_5+3_corrected	TAALSFPAAPIITFPKADIARVPSEESTTLVQNPLLLLENKE--SVEKTPTTIKYFRTBI
mouse_5_corrected	-----
rat	NHVIPTASMTSAPTSVSLGKSPVDNSGHLSPGTTIQTGKDSVETPLPSPLSTP--SIP
human_5+3_corrected	SQVTPPGAVMTYAPTSIPMEKTHKVNASYPRVSSNEAKRDSVITSSLSGATTKPMTII
mouse_5_corrected	-----
rat	TSTKFSKRKTELHQIFVNNQKKEGMLKNFYQFGLQKNPAKLPKIAPELLPTGQSSPSDST
human_5+3_corrected	AITRFSRRKIETWQNFVNNHNPFGRLRNQHKVSLQKSTAVMLPKTSPALPQROSSPFHT
mouse_5_corrected	-----
rat	TILTSFPFALSTTMAATONKGTVEVVGARSLSAGKQ-PFTNSSPVLPSITIKRSNTLNF
human_5+3_corrected	TILSTVMQIPENTLTATHTTTKTHNPG-SLPT-KKELPFPFLNPMPLPSIISKDSSTKSI
mouse_5_corrected	-----
rat	ISTETPT-VTSPTATASVIMSETQRTSKEAKDQIKG-P-RKNRNANTTPROVSGYSAY
human_5+3_corrected	ISTQTAIPATPTTFASVITYETQTERSRQAQTIQEQEPQKNRTDPNISPDQSSGFTTF
mouse_5_corrected	-----
rat	SALTADTPLAFSHSPRODDGNGVSAVAYESTTS--LLAITELFEKYTQTLGNTTAELETT
human_5+3_corrected	TAMTP--PALAFTHSPPENTTIGISSTISFHSRTLNLTDVIEELAQASTQTLKSTIASETT
mouse_5_corrected	-----
rat	LLSKSQESTTVKRAS-DTP-FPLLSSGAPFVPTSPPPFTKGVVTDKVTSAFQMTSNRV
human_5+3_corrected	LSSKSHQSTTRKASLDTEPFPFLSSSATIMPVPISEFFTQRAVTITRGDSHFRLMNTV
mouse_5_corrected	-----
rat	VTIYESSRHNTDLQQPSAEASPNPEIITGTTDSPSNLFFSTSVPAIRVDKPNQSKWKPSF
human_5+3_corrected	VKLHESSRHN-LQMPSSQLEP-----LTSSTSNLLHSTPMPALTTVKSQNSKLTSP
mouse_5_corrected	-----
rat	WFEHKYQLKSYSETIERGKRPAVSMSPHLSLPEASTHASHWNTQKHASKSVFDKPKGQNP
human_5+3_corrected	WAEQFWHKPYSDIAEKGGKFEVSMLATTGLSEATTIVSDWDGQKNTKKSDFDKKPVQEA
mouse_5_corrected	-----

rat -TSKHLPIYVSLPKTLLKKPRIIGGKAASFTVPANSDFLPCAVGDLPIIHWTRVSSGX
human_5+3_corrected TTSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAFLPCEAVGNPLPTIHWTRVS-GL
mouse_5_corrected -----

rat EISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFFVLSVVFYPARYLD
human_5+3_corrected DLSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCASNLFGTDHLHVLSVVSYPRIILE
mouse_5_corrected -----

rat RHVKEITVHFGSTVELKCRVEGMPRETIVSWILANQTVVSETAKGSRKVVVTPDGTLLIYN
human_5+3_corrected RRTKEITVHSGSTVELKCRAEGRPSFTVTWILANQTVVSESSQGSROAVTVTDGTLVLHN
mouse_5_corrected -----

rat LSLYDRGFYKCVASNPSSQDSSLVVKIQVITAPPVITEQKROAIVGVLGGSLLPCTAKGT
human_5+3_corrected LSIYDRGFYKCVASNPSSQDSSLVVKIQVIAAPPVILEQRRQVIVGTWGESLKPCTAKGT
mouse_5_corrected -----

rat PQPSVHWVLYDGTTELKPLQLTHSRFFLYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV
human_5+3_corrected PQPSVHWVLYDGTTEVKPLQFTNSKLLFLSNGTLYIRNLASSDRGTYECIATSSSTGSERRV
mouse_5_corrected -----

rat VILTVEEGETIPRIETASQWTEVNLGEKLLLNCSATGDFKPRIIWRLPKFAVIDQWHRM
human_5+3_corrected VMLTMEERVTSPIEASQKRTVMTGDKLLLNCSATGEPKPOIMWRLPSKAVVDQ-----
mouse_5_corrected -----

rat GSRIHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLVMHVLRILTPAKIEQKQYTKKQ
human_5+3_corrected GSWIHVYPNGSLFIGSVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDHKQYFRKQ
mouse_5_corrected -----

rat VLHGKDFQVDCRASGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTTLFNGTLYFNKV
human_5+3_corrected VLHGKDFQVDCRASGSPVPEISWSLPGDTNINNVAQADDSGHRTRRYTTLFNGTLYFNKV
mouse_5_corrected -----

rat GMAEEGDYICSAQNTLKGDEMVKHLTVLTAIPRIQSYKTMLRAGETAVIDCEVTGEF
human_5+3_corrected GVAEEGDYTCYAQNTLKGDEMVKHLTVITAAPRIQSNKTNKRIKAGDTAVIDCEVTGDF
mouse_5_corrected -----

rat KPNVFWLLPSNNVISFSNDRFTFRANRLSIHKVKPLDSGDYVCVAQNPSGDDTKYKLD
human_5+3_corrected KPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYKLD
mouse_5_corrected -----

rat IVSKPPLINGLYANKTVIKATAIRHSKRYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSR
human_5+3_corrected VVSKPPLINGLYTNRTVIKATAVRHSKXHFDCRAEGTPSPEVMWIMPDNIFLTAPYFGSR
mouse_5_corrected -----

rat VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVQLEVLEMLRRPTFRNPFNEKIVA
human_5+3_corrected ITVHKNGTLEIRNVRLSDSADFTCVARNEGGESVLVQLEVLEMLRRPTFRNPFNEKIVA
mouse_5_corrected -----

rat QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPHNSPYTMAGNGSLILYKATRNKSGKY
human_5+3_corrected QLKGSTALNCSVDGNPPPEIIWILPNGTRFSNGPQSYQYLIASNGSFII SKTTREDAGKY
mouse_5_corrected -----

rat RCAARNKVGYTEKLILLEIGQKPVILTYBPGMVKSVSSEPLSLHCVSDGIPKPNVKWTFP
human_5+3_corrected RCAARNKVGYTEKLVILEIGQKPVILTYAPGTVKGISGSLSLHCVSDGIPKPNIKWTMP
mouse_5_corrected -----

```
rat      GGHVIDRPQVDGKYILHENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVAIFPK
human_5+3_corrected  SGYVVDRPQINGKYILHDNGTLVIKEATAYDRGNVICRAQNSVGHTLITVPVMIVAYPPR
mouse_5_corrected  -----
```

```
rat      IINYLFRNMLRRTGEANQLHCVALGIPKPKVTWETFRHSLLSKATARKPRSEMLHPQGT
human_5+3_corrected  ITNRPPRSIVTRTGAAPQLHCVALGVFKPEITWEMPDHSILLSTASKERTHGSEQLHLQGT
mouse_5_corrected  -----
```

```
rat      LVIQNLQTSDSGVYKCRQNLGTDYATTYIQV
human_5+3_corrected  LVIQNPQTSDSGLYKCTAKNPLGSDYAATTYIQV
mouse_5_corrected  -----
```

(rat: SEQ ID NO:10)

(human_5+3_corrected: SEQ ID NO:11)

(mouse_5_corrected: SEQ ID NO:12)

•

[illegible]

rat
 human_5+3_corrected
 PITGVSRRPSSSDISSHTTADPSFSSHPGSGHTTASSLPHIPRNNTGNEFLSRHLGRERT
 SVREVSEPRHNFYSHTTQILSTSTFSPDPHTAAHSQFFIPRNS-TVNIPLFRFPGRQRK
 .: **,* ..: **** * *;:***:** * * ****, * *:* *;:***:*.

rat
 human_5+3_corrected
 IWSRGVKNPHRTFVLRHRHRTVRLPAIKGFANKNVSOVPEATEYPMCHTCPSAEGLTVA
 IGGRGRIISPYRTFVLRHRYSIIFRSTTRGSSEKSTAFSATVINVLTCLSLPRERLTTA
 * ,***: .*:*****: .*: :*: :*: :*. ** * : * * * *

rat
 human_5+3_corrected
 TAALSVPSSSHSALEPKTMNVGVIAEESTTVVKPLLLFKDKQNVDIRIITTTKYSGGES
 TAALSFPSAAPITFPKADIARVPSEESTTLVQNPLLLLENKP--SVEKTTPTIKYFRTET
 *****,: :***: . * ;*****:***:***:***: . : * * * * *

rat
 human_5+3_corrected
 NHVIPTASMTSAFTSVSLGKSPVDNSGHLSPGTIOTGKDSVETTLPSPSLTP--SIP
 SQVTPTGAVMTYAPTISIEMKTHKVNASYPRVSTNEAKRDSVITSSLSGAIKPTNTII
 .: * * * * * ****: .: * : * : : * : : * : : * : : * : : * : : * : : * :

rat
 human_5+3_corrected
 TSTKFSKRKTPHLQIFVNNQKKEGMLKNPYQFGLQKNPAKLPKIAFLPTGQSSPSDST
 AITRFSRRKIPWQNFVNNHNPCKGRLANQHKVSLQKSTAVMLPATSPALPQKQSSPFHFT
 : * : * : * * * * * : * : * : * : * : * : * : * : * : * : * : * : *

rat
 human_5+3_corrected
 TLLTSPPPALSTTMAATQNGTEVVSGARLSAGKKOPFTNSSPVLPSTISKRSNTLNLFL
 TLSTSVMQIPSNLTHTAHTTTKTHNPG-SLPTKKELPPPLNPMPLSIISKDSSTKSII
 ** * * * * : * : : : : * : . . . * : : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 STETPT-VTSPTATASVIMSETQRTSRKEAKDQIKG-P-RKNRMNANTTPROVSGYSAYS
 STQTAIPATPTTFPASVITYETQTERSRAQTIQREOEPQKNRTDPNLSFDQSSGFTIPT
 ***: * , : * : * * * * * * * * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 ALTADTPLAFSHSPRODDGNGVSAVAYHSTTS--LHATLFEKYTQILGNTTALETTL
 AMTP--PALAFTHSPPENTTGISSTISFHSRTLNLTVDVIEELAQASTQTLKSTIASETTL
 * : * , . * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 LSKSQESTTVKRAS-DTP-PPLLSSGAPPVPTSPPPFTKGVVTDKSVTSAFQMTSNRVV
 SSKSHQSTTTRKASLDTPIPPFLSSSATLMPVPISEPPFTQRAVTDTRGDSHFLMTNTVV
 ***: : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 TIYESSRHNTDLOQPSAEASPNPEIITGTTDSPSNLFPSTSVPAKLVKDPQNSKWKPSFW
 KLHESSRH--LQMPSSOLEP-----LTSSTSNLLHSTMPALTTVKSQNSKLTSPSPW
 .: : * : * : * * * : * , * , * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 PEHKYQLKSYSETIEKGRPAVMSPHLSLPEASTEASHWNTQKHAESVFDKPKGQNP-
 AEYQFHWKPYSDIAEKGGKPEVSMLATTLGLSEATTIVSDWDGQKNTKKSDFKPKVQEAT
 ,* : : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 TSKHLPYVSLPKTLLKKPRIIGGKAASFTVPANSDFVLPCEAVGDPLPIIHWTRVSSGXE
 TSKLLPFDSLSRYIFEKRIYGGKAASFTIPANSDFVLPCEAVGNPLPIIHWTRVS-GLD
 *** * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 ISQGTQKSRFHVLPNGTLSIQRVSIQDRGOYLCSAFNPLGVDHFVSLSVVFPARILDR
 LSRGNQNSRVQVLPNGTLSIQRVEIQDRGOYLCSANLFGTDHLHVTLVSVSYPPRIER
 : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 HVKEITVHFGSTVELKCRVEGMFRPTVSWILANQTVVSETAKGSRKVWVTPDGLTLLIYNL
 RTKEITVHSGSTVELKCRAGRPSPPTVWILANQTVVSESSQGSROAVTVVDGLTVLHNL
 : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 SLYDRGFYKCVASNPQSDSLLVKIQVITAPPVILEQKROAIVGVLGGSILKLPCTAKGTP
 SIYDRGFYKCVASNPQSDSLLVKIQVIAAPPVILEQKROAIVGVLGGSILKLPCTAKGTP
 * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 QPSVHWVLYDGTLEKPLQTHSRFFLYPNTLYIRSIAPSVRGTYECIATSSSGSERRRV
 QPSVHWVLYDGTLEKPLQTHSRFFLYPNTLYIRSIAPSVRGTYECIATSSSGSERRRV
 ***: * : * : * : * : * : * : * : * : * : * : * : * : * : * : * :

rat
 human_5+3_corrected
 ILTVEEGTIPRIETASQKNTVNLGEKLLNCSATGDPKPRIIWRLPKAVIDQNHMRMG
 MLTMEERVTSPIREASQKRTVNFQDKLLNCSATGEPKQIHWRLPSKAVVDQ----G
 : * : * * * * * : * : * : * : * : * : * : * : * : * : * : * : * :

(rat: SEQ ID NO:13)
(human 5+3 corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
 VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
 KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
 KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGNPWTCDCHLKWLSEWMQGNP
 (SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPPNVE
 RINLGYNLSLVRLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
 LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
 QIFKISFIKFLYLSDNFLTSLPQEMSYMPDLDSLLYLHGNPWTCDCHLRWLSDWIQPDVI
 KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
 AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
 VCNIDYGHIPVWQILALYSDSPLILERSHLLSETPQLYYKYKVAPKPEDIFTNIEADLR
 ADPSWLMQDQISLQLNRTATTFTSLQIQYSSDAQITLPRAEVRPVKHKWTMISRDNNTK
 LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
 ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGIP
 DASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYRCVAANPSGVDFLIFQV
 SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRH
 NYRELTQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALEKAKKNAMPDKRENTTV
 SPPPVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
 YGTEFSPVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG
 ATEFQDSDQMGRGREHFQSRPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
 SEPRHNHFYSHTTQILSTSTFPSPDHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS
 PYRTPVLRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSPAAPI
 TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
 SIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLGAIKPPMTIIAITRFSRRKIPWQQNFV
 NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNTLTT
 AHHTTTKTHNPGSLPTKKELFPPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
 TQTERSRAQTIQREQEPQKKNRTPDNISPDQSSGFTTPTAMTPPALAFTHSPPEENTTGISST
 ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETTLSSKSHQSTTTRKASLDTPIPFLSSSAT
 LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLMQMPSSQLEPLTSSTSNLL
 HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGGKPEVSMLATTGLSEATTLV
 SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAF
 LPCEAVGNPLPTIHWTRVSGDLDSRGNQNSRVQVLPNGTSLIQRVEIQDRFQYLCSASNL
 FGTDHLHVTLVSVSYPRIERRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSE
 SSQGSRQAVVTVDGTLVLHNLISYDRGFYKCVASNPGGQDSSLVKIQVIAAPPVILEQRR
 QVIVGTWGESLKLPCAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPIREAASQKRTEVNFGDKLLLNC SATGEP
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTN
 KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVSKPPLINGLYTNRTVIKATAVRHSHKHFDCRAEGTP
 SPEVMWIMPDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGEVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEIHWILPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVG YIEKLVILEIGQKPVILTYAPGTVKGISGE
 SLSLHCVSDGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGNYICKA
 QNSVGHTLITVPVMIVA YPPRITNRPPRSIVTRTGAA FQLHCV ALGV PKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGTLVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

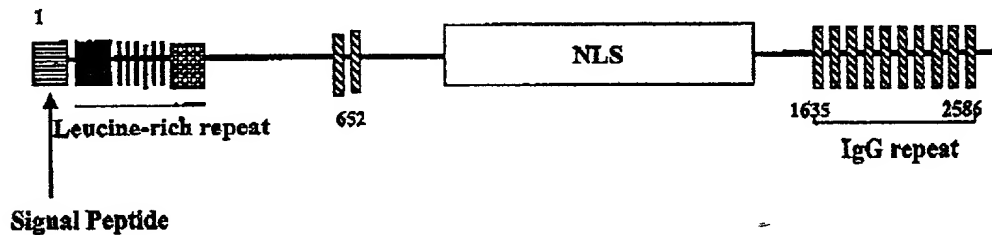


Figure 20

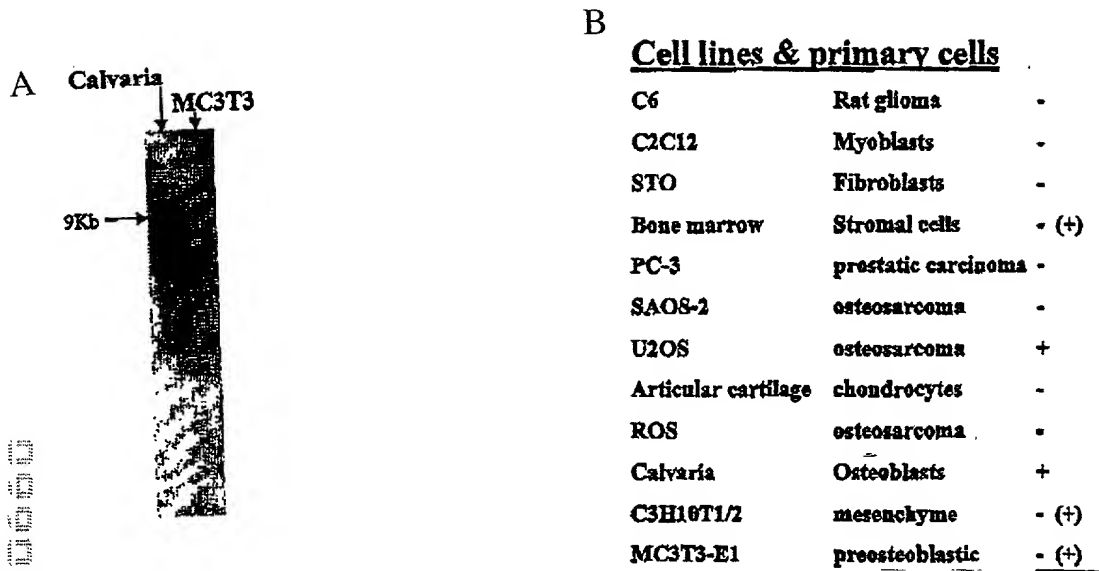
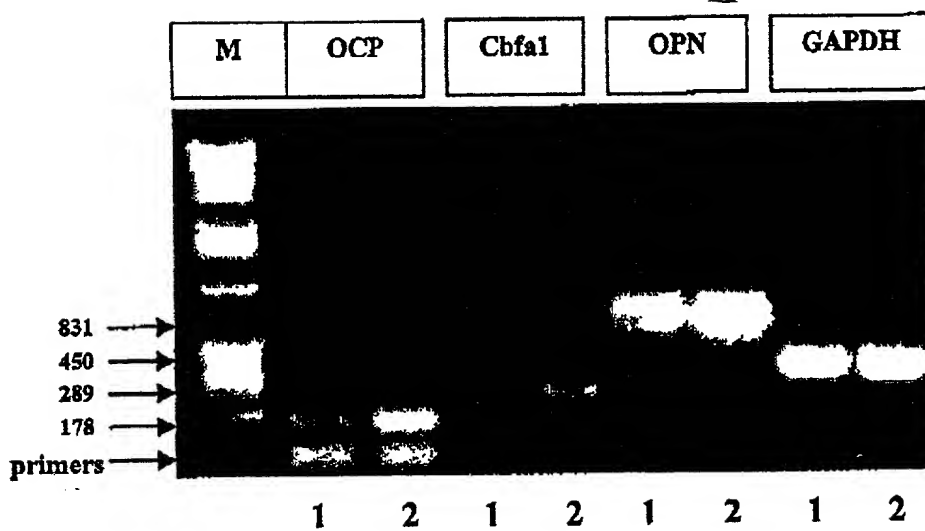


Figure 21



52/85

Figure 24

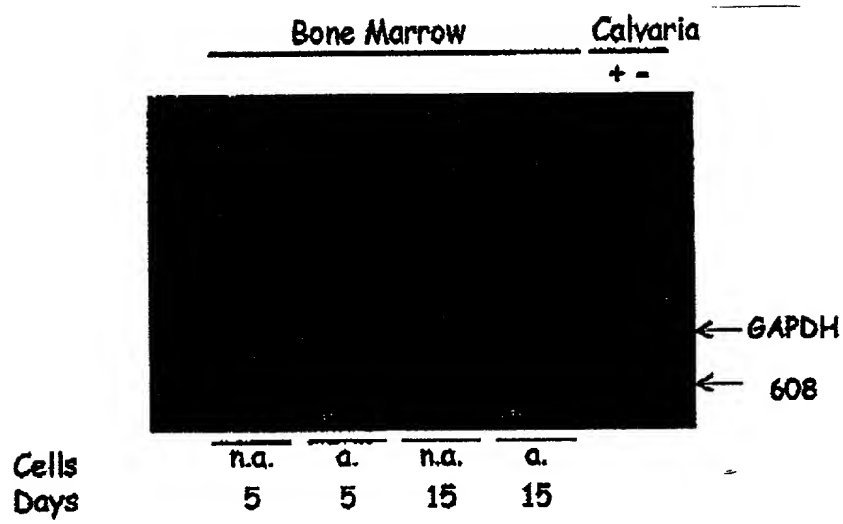
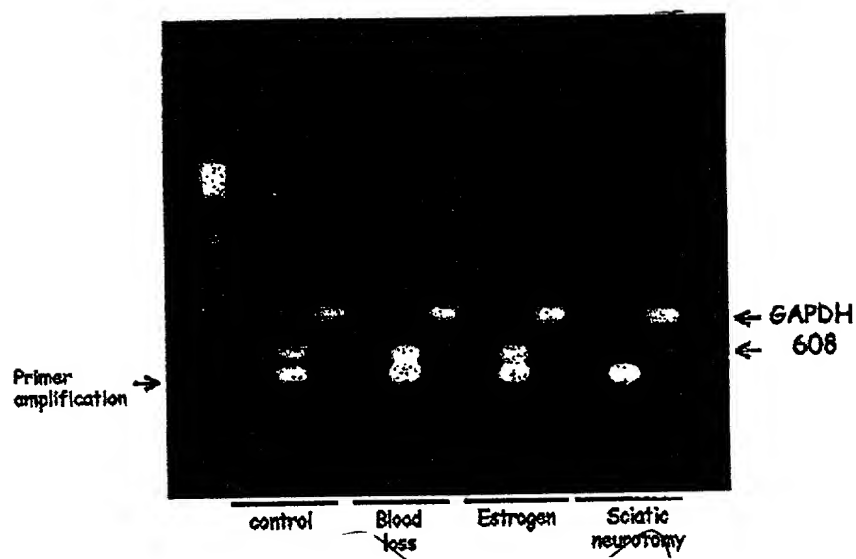


Figure 25



[illegible]

A high-contrast, black and white photograph of a person standing, wearing a dark, textured garment. The image is heavily degraded with significant noise and artifacts, including a large white vertical streak down the center and a small circular mark near the bottom left.

Figure 27

A

B

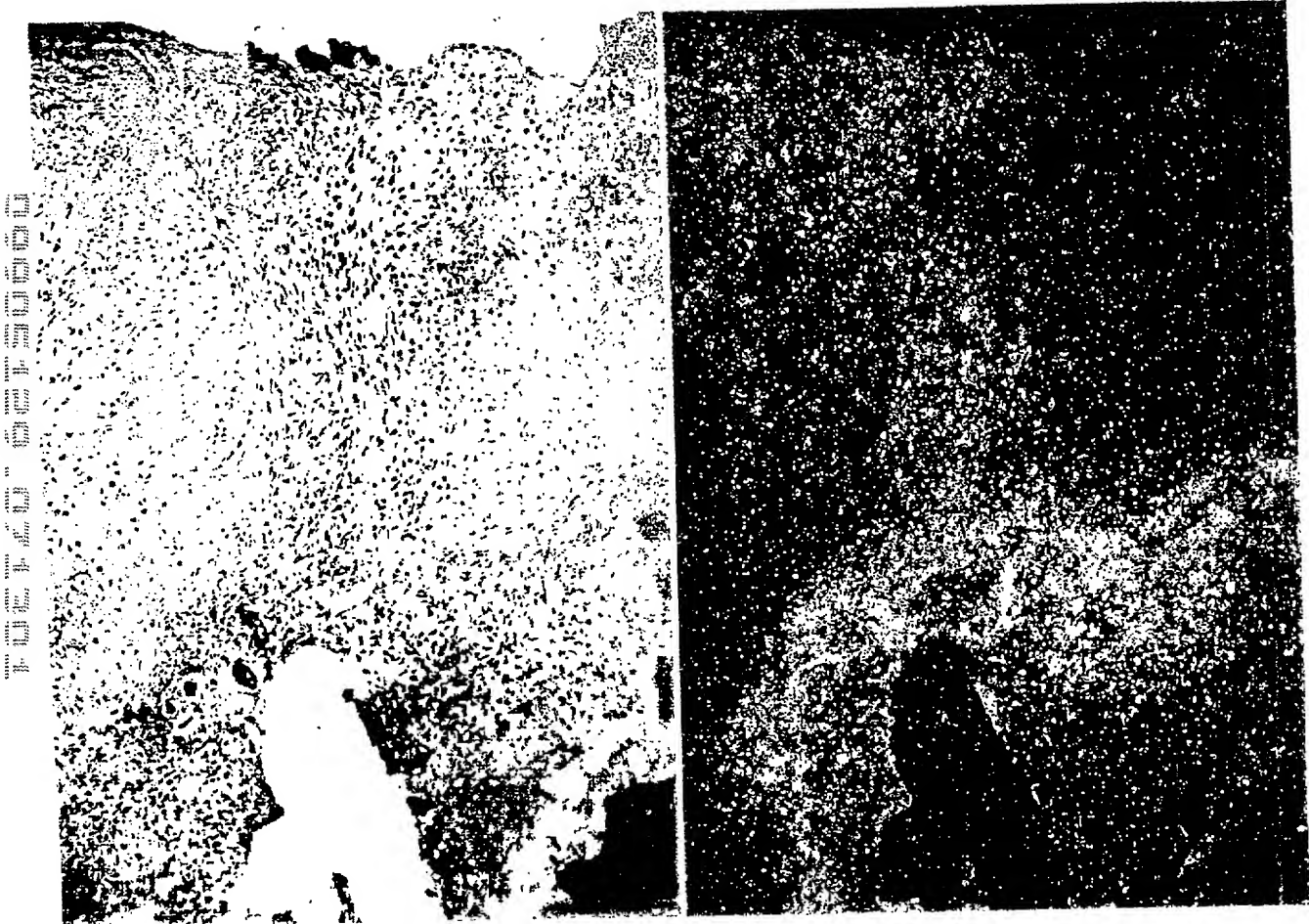
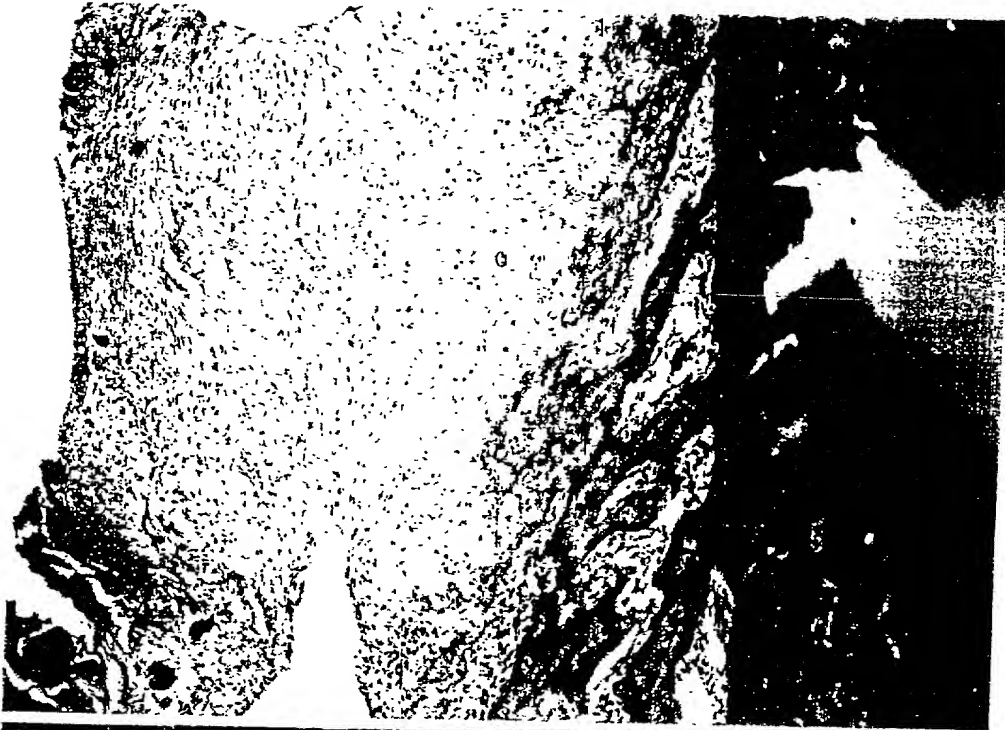


Figure 28

A

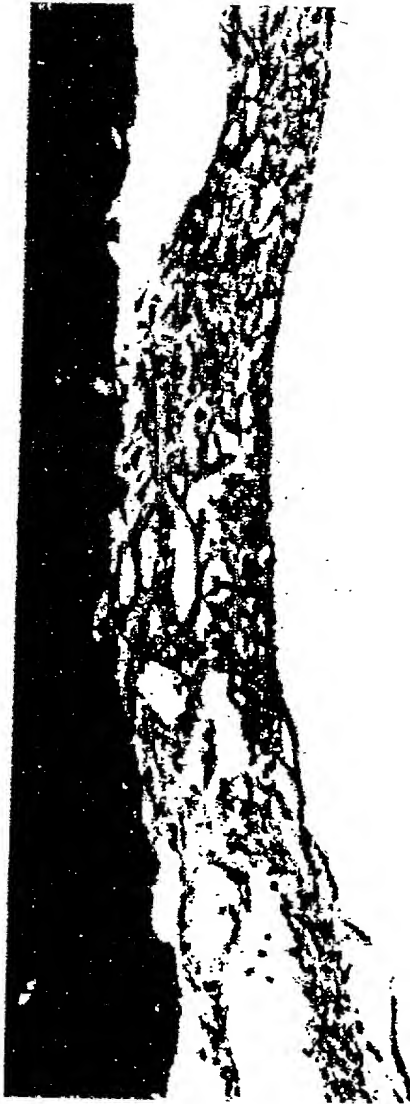


B

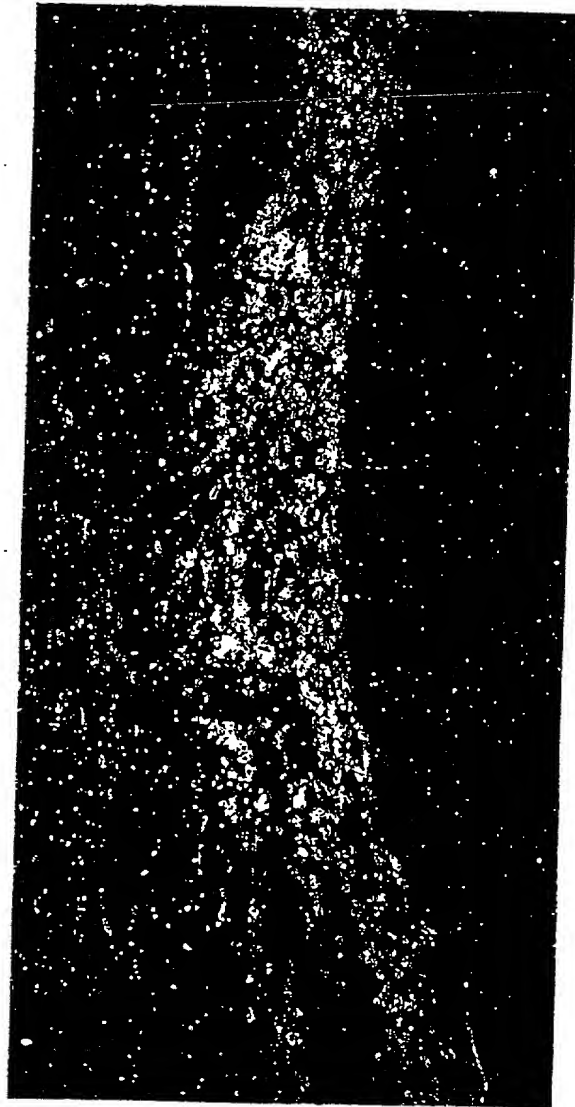


Figure 29

A

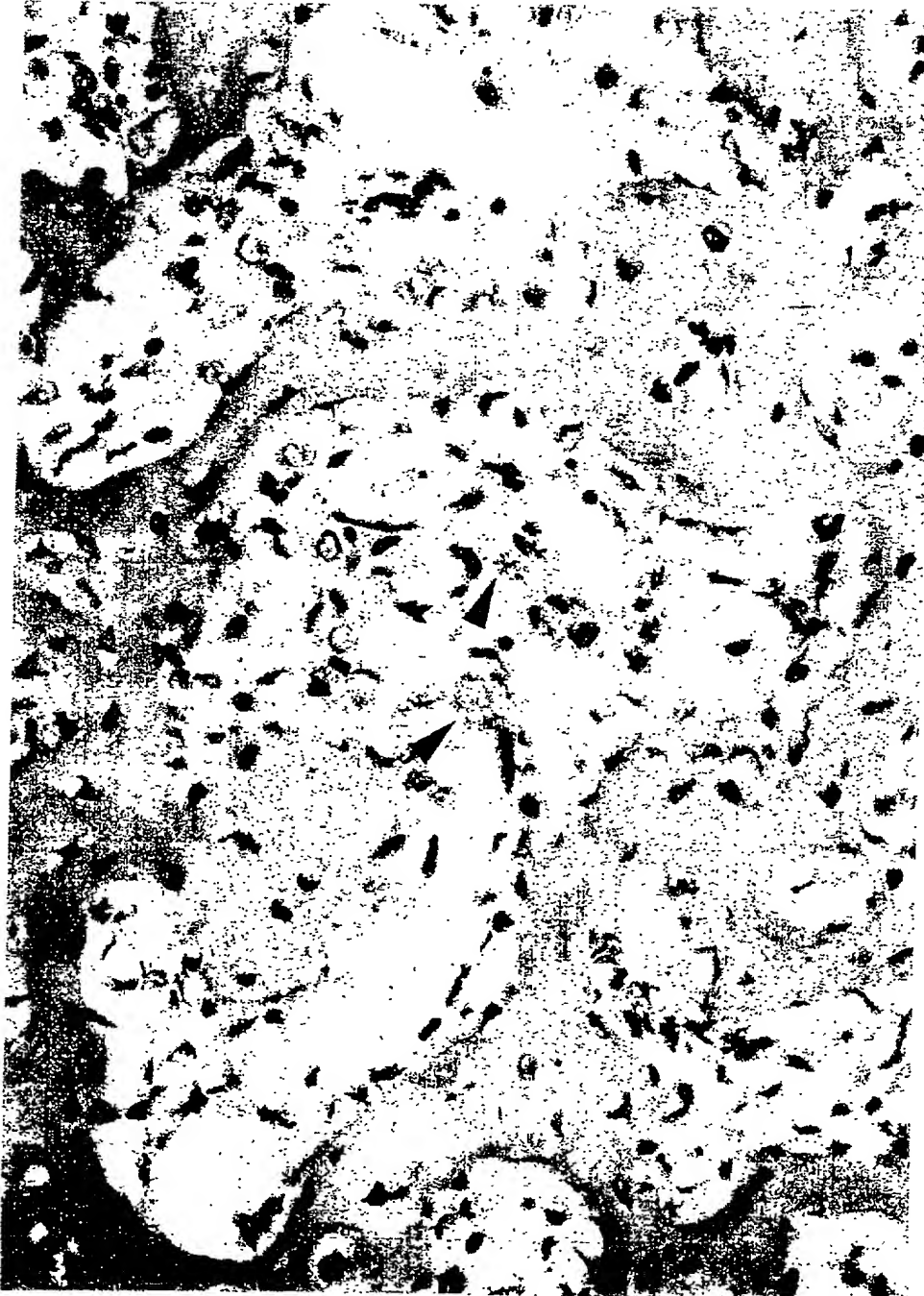


B



57/85

Figure 30

[illegible]

58/85

Figure 31



2025 RELEASE UNDER E.O. 14176

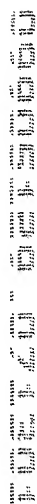
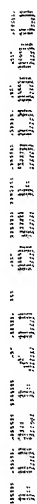
[illegible][illegible][illegible]

Figure 33

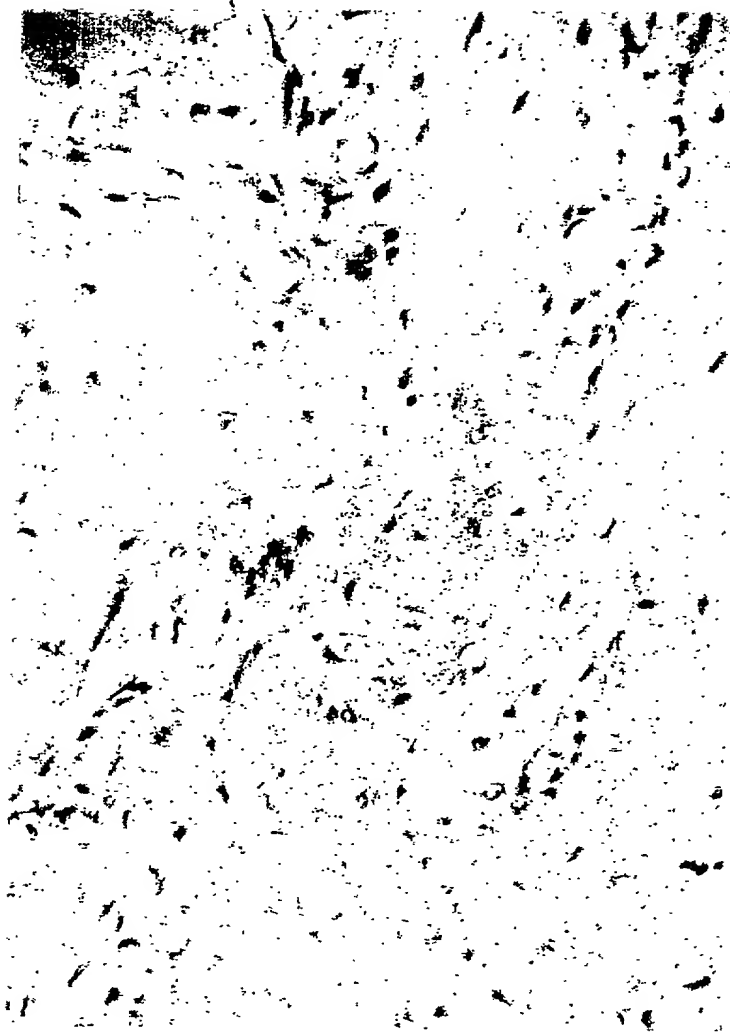


Figure 34

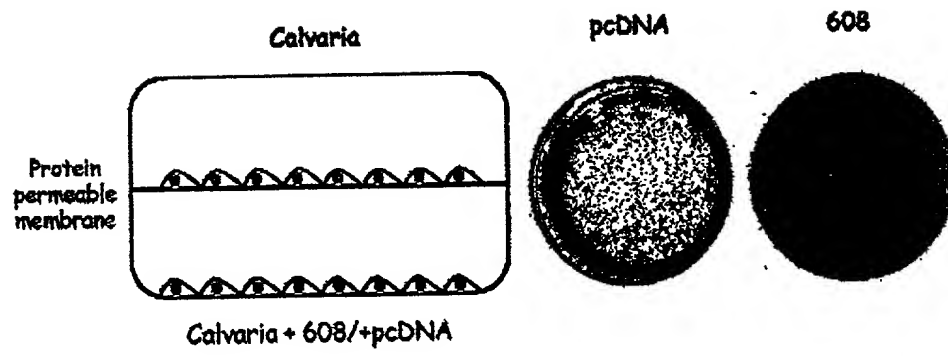


Figure 35

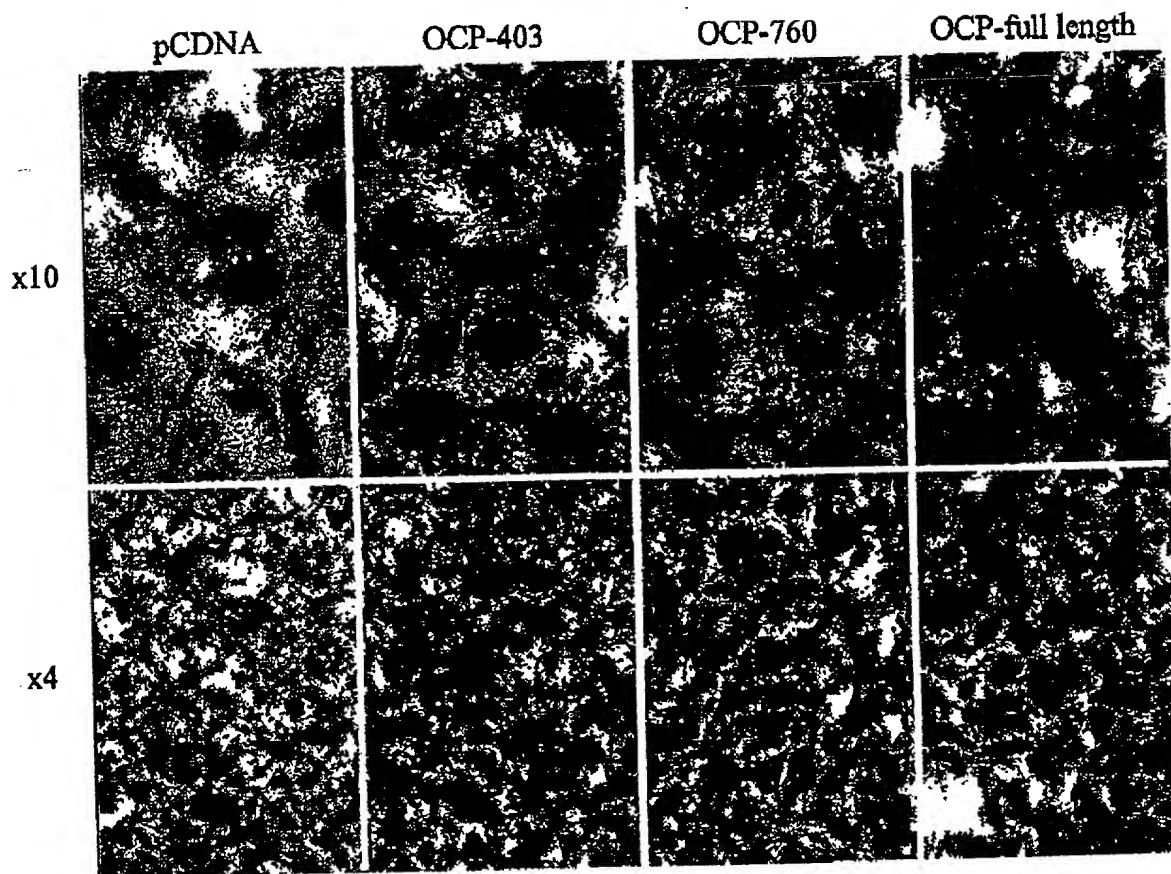


Figure 36

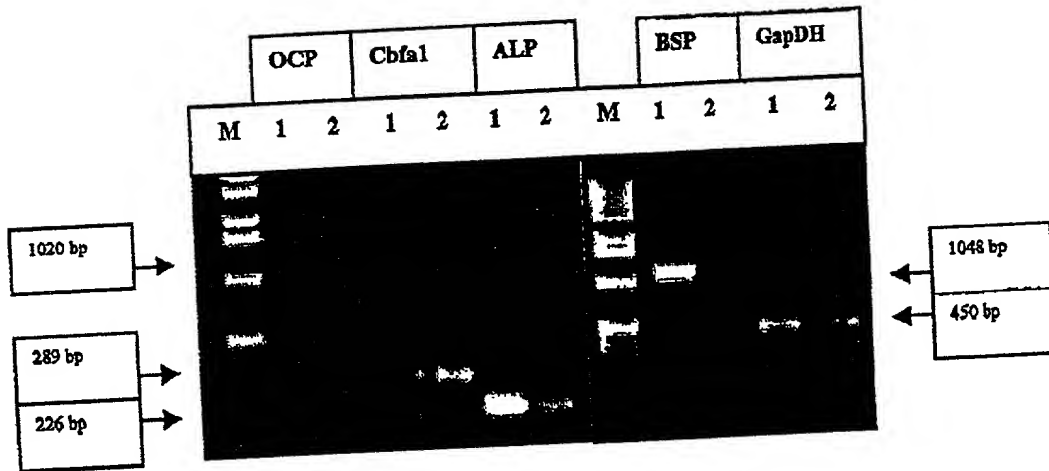


Figure 37

x 4

x 10

pCDNA
ROS stable line

OCP
ROS stable line

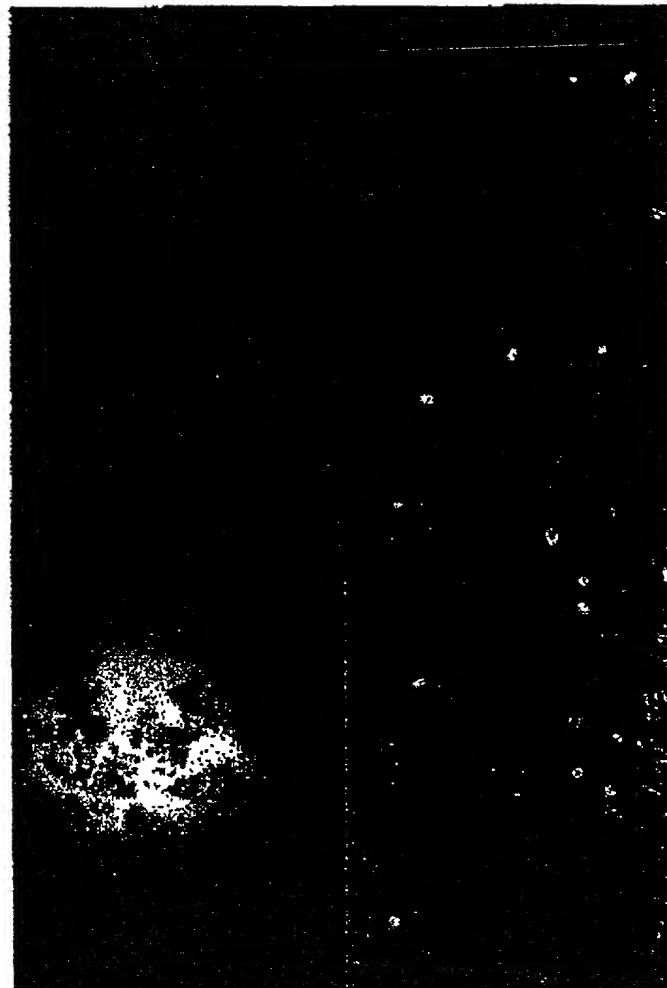


Figure 38

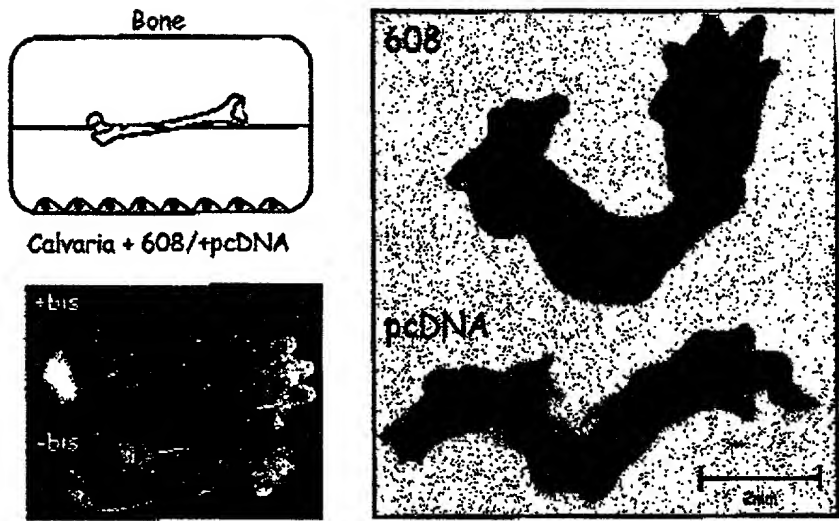


Figure 39

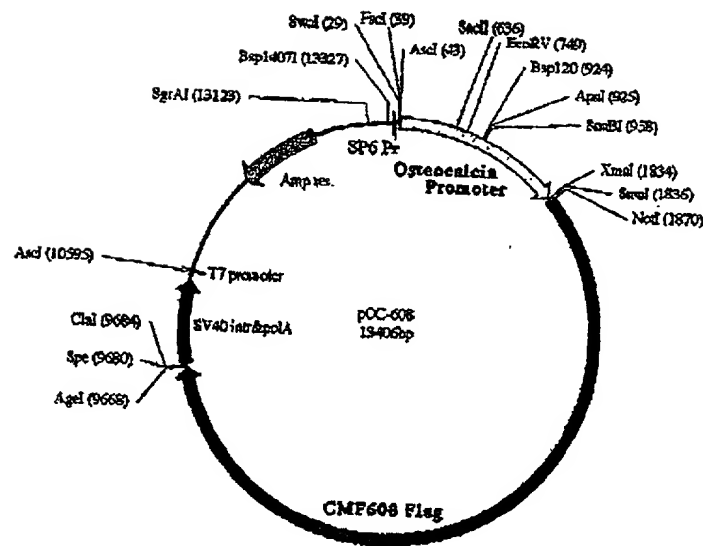


Figure 40

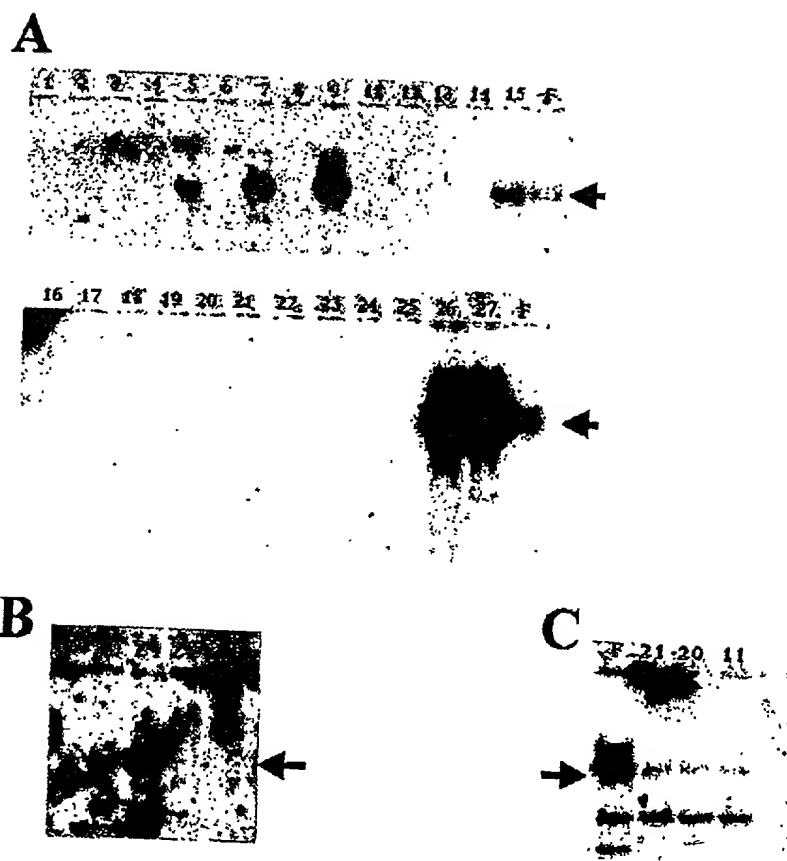


Figure 41

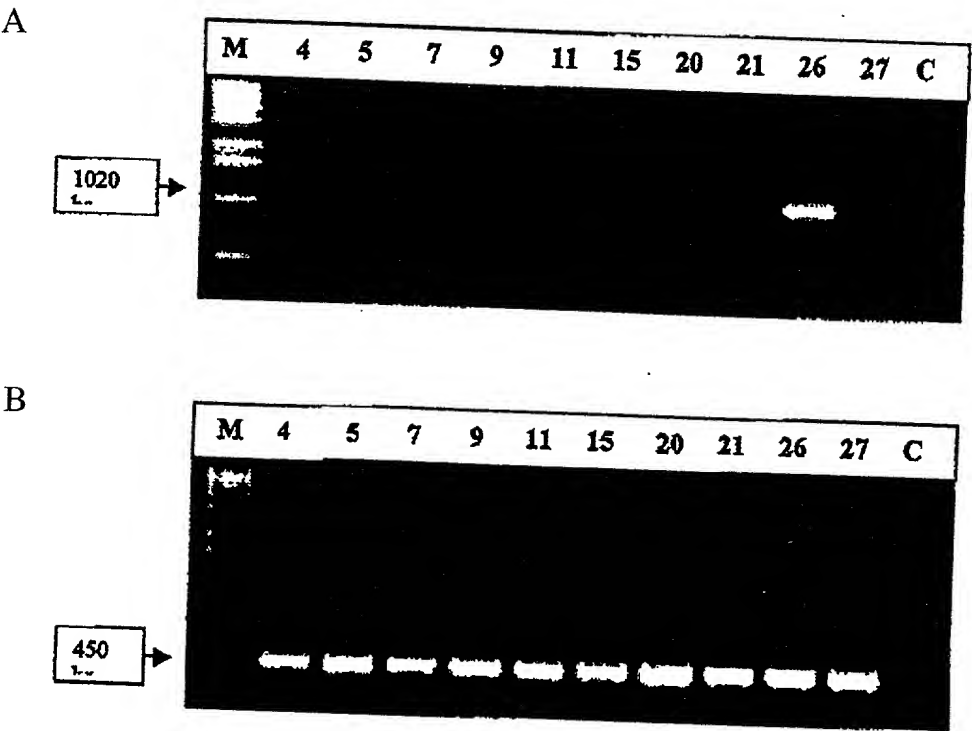
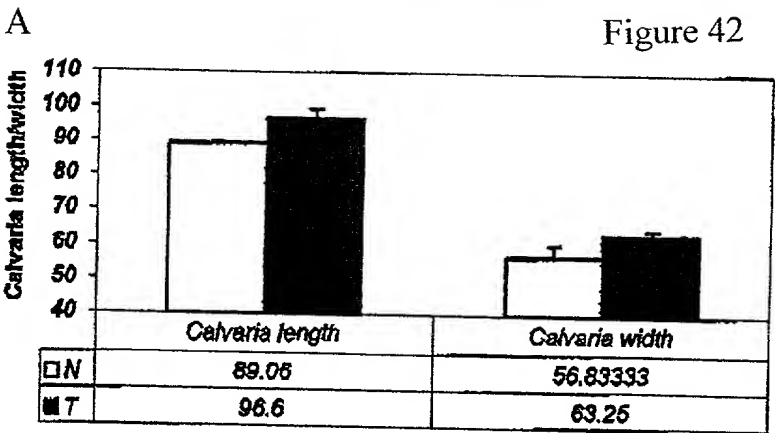


Figure 42



B

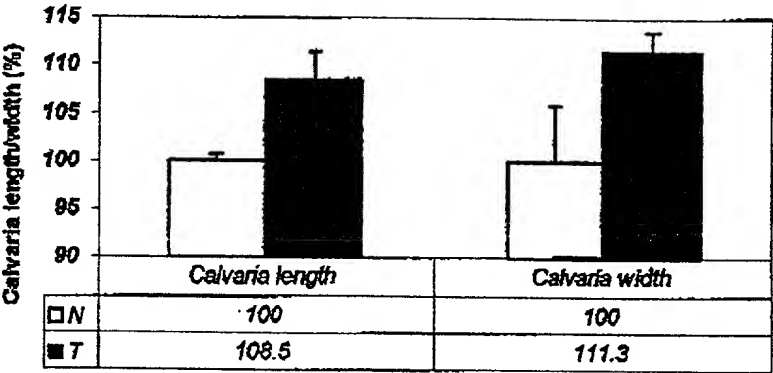


Figure 43

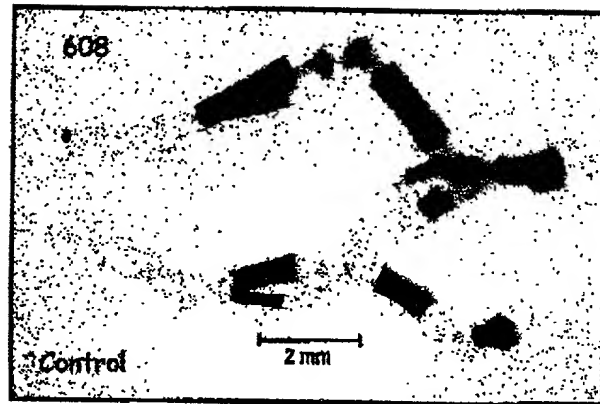


Figure 44

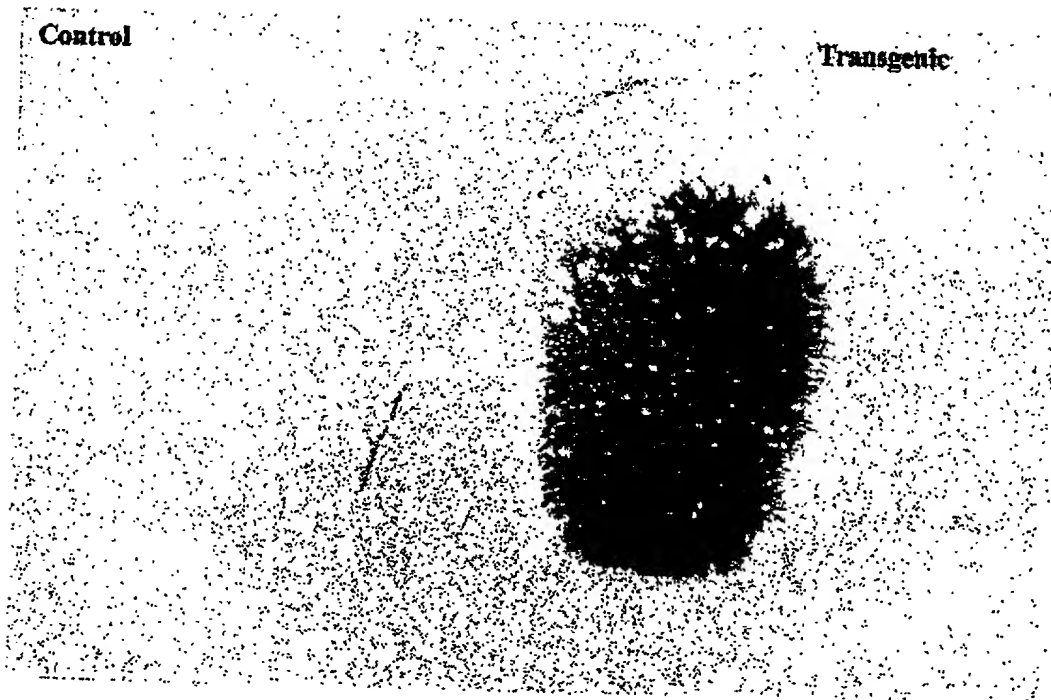
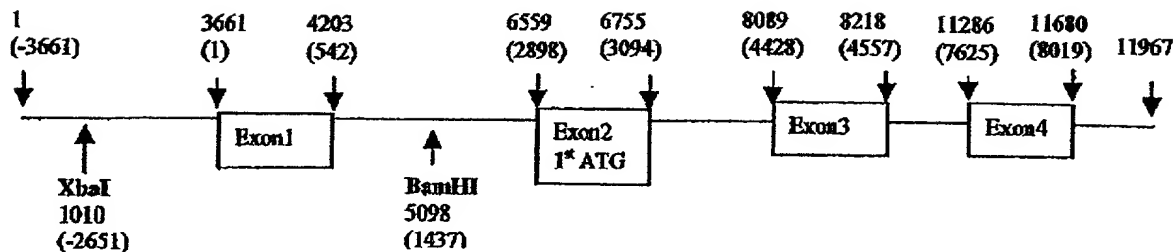


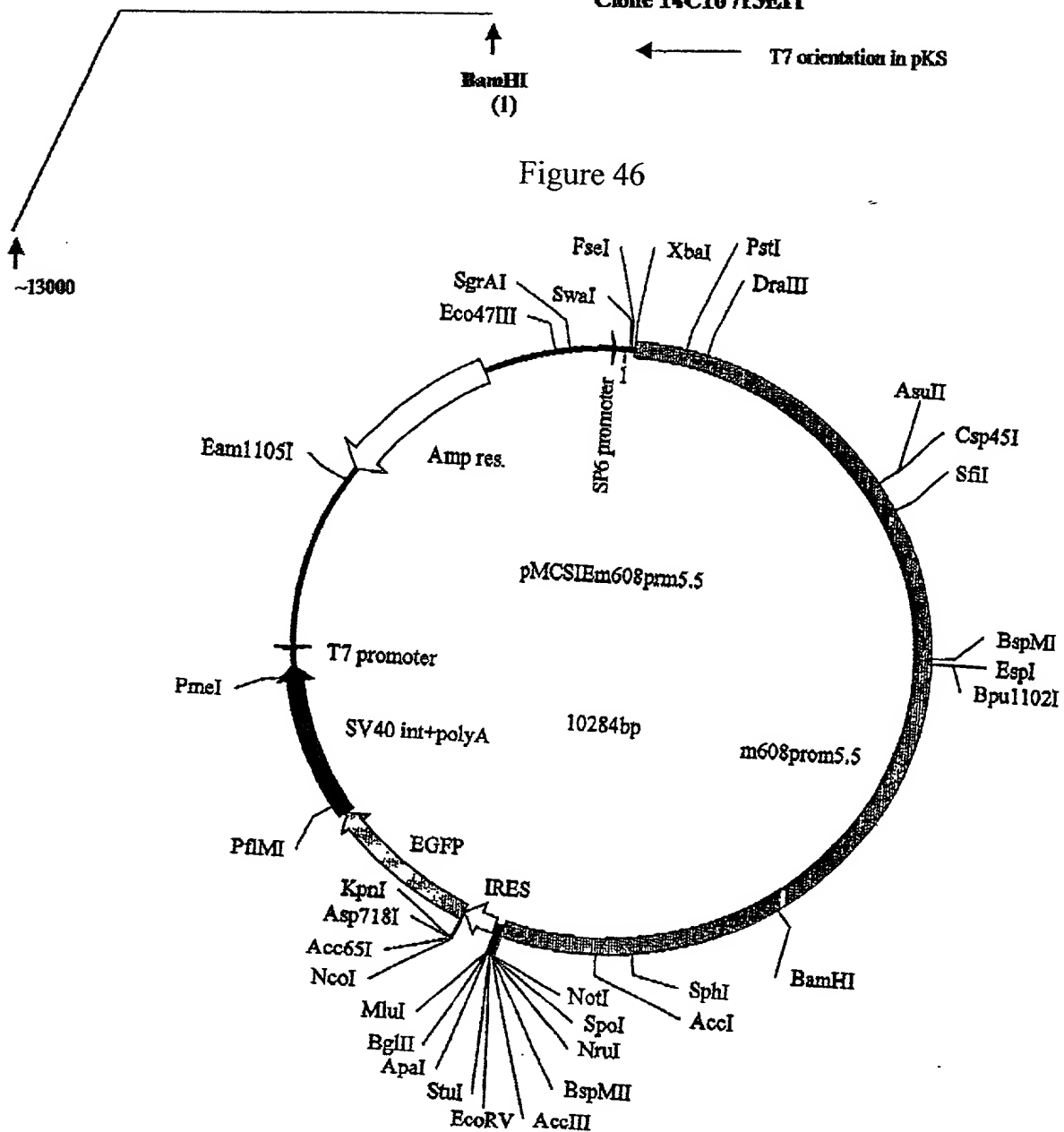
Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46



↓ (XbaI)

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAAG
GGTCCAGGAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA
CATCGCTTCAGTGACTGAAGCCAGGCCAAAAGGAGCGGGGAAGGATTATATGTAGCCTGGGA
CGCTTTCATAAACACTGATGACGTGTTTGTCGAAAGCAAGCAATTTGAGGAGAAAACGGCT
GGGACGTCCGAAAGAAGGAGTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAAC
TAACATAACCTATACTATTGAGGGAGAAAGGAAGGACATTCAGCAGCAGCAGCAGCAGCA
GCAATCAGATAAAGGAAGGCTTTGGTTAGTTTGGAATGTATGATACCATTTAAATAACA
GAAAGCGCCTCCAGTTCTCTGAAGAGTCAAGTCCCCAGCTAGTGAAGACTAAGCCTACTAA
GCCTTTTGCTCCCGTTGGAAGCAAAGCTTCCTTCAATCAGGTGAAGGCTCTCTCTCAGA

AGATTTCCTGCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGG
 TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACATCATGTGGTA
 TTCATTTACGGCTCTCTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
 CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
 AOCGAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACAGTGTAGGCATAGAAGTA
 ATCATTAGTAAATCTTAATTTGGGTAAACTCATTACATAACAGCTCCAGGTTGGGAGGGAT
 CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCCTC
 CTTGGCCATGCTCCCATCACTGTGTCAAGTAAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAG
 AGCAGACAGTGTATGCCTGCAAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACT
 GTCCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGC
 AGCAGCTAGAGAAAGAAATTTGACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTAT
 TTTGGAAGGTAAAGCCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAAATT
 GATCATATAAAGAAGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTATCTAT
 TTAAGTGTGCAAAAGTTTCTAAATTCACATGTACTATTTTATATTTATAAAAAATATTTT
 AAAATGTATATAAAAGTGTAAAAGGCTCTTGGTCAAAACAAGAGAGTTAAATTTACAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
 AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCAATG
 GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAAGTTTACA
 CTCTAGGGAATAAGAGTTCTCTTTTCCAGATTTCACTAGCATTGTTGTGCATCATTTATC
 TTCTTGATGATGAGCATTATAAGTGGAAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
 ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCAATTTATTGGATAA
 TTGGGGGATGGTGTTAATTTTGTGAGTTCTTATGGAATTCCAAAAACAAAAACAAAC
 AAACAAACAAAAACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
 AGAAATCTGCACAGAATTTATTGCTACATTGTTTATTACGACAGCCAAGAATGTGGA
 ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAATGTGTACAAC
 AGAGTCCCATGTGGCCATAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
 GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAGGAAACACCGTGTAGCTT
 CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATT
 TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
 AGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
 GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
 GGCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
 TTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAGGGGCGTTCCAGTTTT
 CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGG
 TTTTGTGCTATTTTTCATTTTCTTCTTACTGTTTGTGTTGTGTTGTTGTTGTTGTTGTTG
 GCTGTGTTTTCATTTGTTTCTTTTGTAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
 AGCTGGAACGATCTGGAAGAAGTTGGGGAAGAGAAAAATTGTATGGAGCATATTTAA
 ACAAAACAAACAAACAAACAAAGGTTTATTTGCCACAAAAAGGTGTGAATTAATTAAC
 JAGTTACGACTCTTAAAGAAAAATATCCCAATTATCCAGAGTTGCTATGTATGCTGTGC
 TAGGACTTTGCTTGAAGTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial

NotI site

(SEQ ID NO:17)

Figure 48

CACAGACCTTCTCTTCTAACCTCTCTCCCCATCTTGTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAACCTGCTCCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTCTGACCACATCCCCAAGTTATCACAAGGAGTTT
TTCTCTCAAACCTTTCTCTCCAGCAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCT
TCATTGGGAACCCCTGTGCTCCATCCAATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT
CTTGGTCAGTGCACTCTTTTGTATCCTCAAGAACACTGGGTCTGAAAATTTTAACCCAAAG
AACTGTTTTTTTGTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTTGTCTCCTCCTTTTCAGTACCACCCCTGCCTTTGTCTGCTTTCTCAAAGAATCAATAAA
ACCAAAGTGATATAGATTTCATGGCATTCTCTAAGTCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAGCCTGGAAGCAGTCACGAGGCAGTTTCACAGAAACTTAGCCCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATGCCCCAGATTGTCCCTGAGATATTGTGGTG
GGTCTTGCACTGCTTTTCTGCACTATTTTACTGGATAAGAGTTAGAAAATCTCAGGGCGAGCT
TAGCAAAAAGTATACCTAGAACTCTTCATGACAGTCAGGTATTGCAAACACTACATTGCATATTA
GAAGAAAGTTGGTAAATTTCTCTGACAAATGGAGATTCCCTACAGATAACTTTAAAGAAC
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAGGGGGGGGGGTACT
GCTTTATTTCATGATAAGGTCGTAGAGCAGAACCCCTGCTGCTAGCTTTTACAAGGTTTCA
AAAGGTGTAGCATAAATTTGTGACTAGAGTGTGAAAATCTTTACCTGTCATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAACATTACCTGGTTCTGTCAGTCCTTTGAAGGCAT
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCCTGCTGTGACACCCCTACTCCTTTGT
TTTCTGTATTATATAAGCCTGGTGTTCCTTTGTGAAAATTACATCCAGATACAGCACTC
CCTTGTGTCTGTGCTCTTTTGTCAATTTCTGGCCAACCTCATGCCACCTGCCAGAACCCCTA
GTCTTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTCCATGGCATCTAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAACCCCATCTTCTCCTCGTTTGCTAATATAGCAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTTCTGTATGTG
TGTCCATTGTTTCTTCAGTTCTGTTGCCCATTAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTCCGTTTCAACGTTTTGCCCTG
TAAACAAGATATATGGTGGCAGTTTATAATCATAATAAGCTTTGAAATAATATATAACTC
CATTCTCATTCTGCTTCCACGCTGAGCATCCTGTTTCCCAGGGACCACAAGAGCATTGTA
AAAGTAGTGATTATGACCTGCTTTGTTCTGTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACCTTGGGGTGACCTGTATCTGTGTTCTGGGTGCTGCTCACTCATA
TTTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCAACATTCTACTTCTGTAAATCTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTCTTCCAACCTCCAAATATTTCAAATCTTTTAGCCCTCAGCCTGGGATTTCTTG
ACTAAGTCTATTGATTTGGAAGATCTCAGTGGTTAGGATTTGCAGTCATGATGTTTCATAOG
TCAGGCTAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCAATGTCAATGTGCTGGGAATGTG
AGTGTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATTCTAGTTTTGAATAATGTTCCCTAGTATACAGCAATTTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTCATATTCTAAGGAT
AGGTCTCATTCAACCACATGTTTATCATTTGGGATCTGCATTTTCTGATGCAAAATGATTT
ATTCTTCCAGAGCACTGGAATTGGGTTGAATCATCTTATAACGGCCAAAACTAAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA
TGACAGAATGTGAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCTTGGCAGGATTCTC
TTGGAACATCACACAAAGAAATTCCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTGAGACCTTGACTTTGAG
GTGAAGGAAAGCTCGCCAACTCCTTTTTATATGTAACACAATATATCAAGATCTAATGTGA
GACAGTATGCCAGTCCCAAGATCTGTCAATATGACTGAAGACACATTGCGATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCAGTTCACATATAAAGCTTTATGAGAAATT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTCTTCATTACATTTT
AAATGCTATCCCAAAAGTCCCTATACTCCCCCTCACCTGCTCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCATATGATCCACCAAGCTTTACATGCTC
ACTATCTGGTCTATTGCAAGAATGGCTGCCGAGCTGATGCAGTCAGATACAGACACCTAC
AGCCAAACAGTGGAAGGAACCTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG

CCCCGGATGGGGAAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
ACCTTGGGGCTCTCAGAGTCTGAACCACCAACCATAGAACATTCATGGGCTGTACCCAG
GCCTCTCCACTCATATGTAAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAGTTCTGAGGTGTGGGTGGAGGGTATACTCAGGGAGGCCCTCACTGC
TAGGAAGAGAAGAGGAGGGGGAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAAATAAAAAATAATAAAAAAAAAAAAAGAATGGCTGCTGAGCCC
TACTCTAAACCATTGCATCCCCCCCCCAATCATTCACTGACTACGAATTAATAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTAACTTCTTTGTGAECACGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTCTTTTCAATTTCCCTGTACATGCATG
TAGCTTAATGATGTTCACTGAATTAATAAATAACAATGAAGAATAAAGACAACCTGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCACCTGGAAGAAATGTCTTCAGTT
TTCTATATATGTTTACTCTATCGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACTTATAAGAAAGTAAACATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAAACATAGAAGTTGAAATTAAGAGTAGACAGGA
GAAGTAATAAGGCAAATAATGAATATTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAAATAAAACATTTCTACTTGAAAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

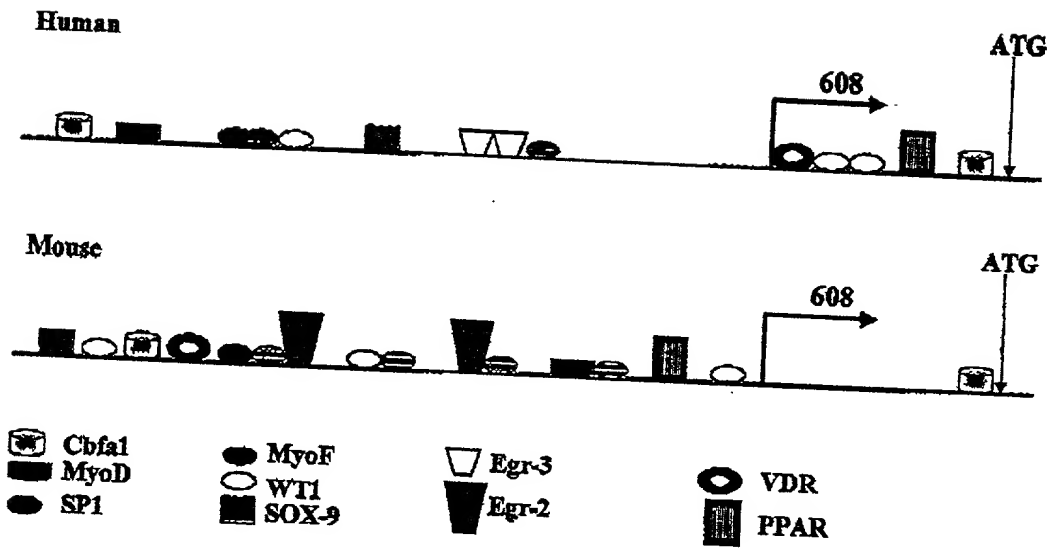


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACCTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTCATAAAATTTT
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGAATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAAACTGAGTCTTAAGGATTCATATTA ACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTAAC TTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCCTAAAGCAATTTAACCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATTTTGAAACATCTAGAAGTAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAGGG
TTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCTGTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTT
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
 TAAACTGACTTTAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
 AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAACCTTTAGAAA
 TTTTGTAAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
 TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
 CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
 AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
 ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
 CAACAAACAACAACGACAACAACCCATAAAAAATTATCACGGCAAATGAACTAAGC
 CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTCTTAATAGTCAA
 ATATGCAGAAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
 GTAGATTAGAGGACACAAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAAT
 GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
 ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTTTT
 ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
 GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
 CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
 GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
 CTTGACTCTAGCTGCATATATATCAAAGATGGCCTAATCGGCCATCACTGGAAAGA
 GAGGCCATTGGACTTGCAAACCTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
 AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
 GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAA
 AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCACTTG
 CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
 AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAACTTA
 CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
 ACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACACAC
 ACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
 TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTC
 TCCTATTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
 AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
 AGGAAAAGCCAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGG
 GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCATTCCCAGAAGCCGAGGCTATA
 GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
 CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
 CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTCAGGG
 CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
 AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGAAGCCAGGCA
 AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
 TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
 TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
 AGGAAAGCTTTGGTTAGTTTGGAATGTATGATACCATTAATAACAGAAGCGCCT
 CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
 TGCTCCCGTTGGAAGCAAAGAACGTTCCCTTCAATCAGGTGAAGGCTCTCCTCAGAAG
 ATTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAG
 GTATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACCTCATG
 TGGTATTCATTTACGGCTCTCTTCTCTTATAACTAAGGTGCATATAGTCTC
 TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
 GCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTTCAGTGATGTAACAGTGACA
 GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTATAA
 CAGCTCCAGGTTGGGAGGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
 TCTAACAAGAGAAGCAGAATTCTTCCTTGGCCATGCTCCCCATCACTGTGTCACTAA
 GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
 GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
 AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
 GACCTTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGGGAAGGTAAAGCCA
 AATAAATAAAACACATATTAAATAATAACAATGTTACAAAAATTGATCATATAAAGA
 AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
 AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATATTTTTAAAATG
 TATATAAAAGTGTAAGGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
 ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
 AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTAGTGATAAAGTCCA
 ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
 GTTTACACTCTAGGGAATAAGAGTTCCCTCTTCTCCAGATTTCACTAGCATTTGTTGT
 CATATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
 GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTCTTTCAGTTCACTAGTC
 TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
 TTCCAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAAACTAGAACTACC
 AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
 TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
 AACGGATAAAGAAAATGTGGAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
 AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
 CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
 CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
 TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
 GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
 ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
 GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
 AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
 ATGATGGATGGTTTTTTGTGCTATTTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTGGT
 GGTGTTGTGTTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
 ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
 AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAGGTTCATTTT
 GCCACAAAAGGTTGTGAATTAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
 AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
 TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
 TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
 CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTCCGGGACCTGA
 CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
 GACCTTGCCCTGATCTCCTTCTCAGAGAGGGGACCACTGATTTTCTTGGTACTTTGCCCC
 CCAAACACCTGTGATTACTTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTT
 ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAAGTGAAGCC
 CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
 AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
 TGTCATTGAATATTTTAAAAAAGCATTTCCTTTGCAATTTATAAATACCCATTACAAA
 ATGGCTTACTTAAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
 GTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
 CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTTCACGTTTGGAG
 GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
 AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
 GGGCTGAGCACTTTGCAATGAGCATTGCTCATTGTGCTGGCAGGTTTTATGATAAC
 TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAAGTCAACTGAGAACATGCCTGA
 AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTCCTTAATTAGTGATTTCATGGG
 GCAGGGCCCAGTCCATTGTTTCGTGGTACCATTTCAGGCACTATTAACAAAAA
 AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
 TGCATCAGCTCCTGCCTCCAGGTTCCCTACCCTATTTGAGTTCCTGTCCTAGCTCCCTA
 CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
 GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
 TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
 AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
 GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
 GTTTGTTTCCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
 ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
 GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
 AGACCTTCTCGGGCTTGACATCCTTGACAGGTGAGATAGGTAGAGGGTGATGGAGGC
 TGAGAAGAGAGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATTCCCGTGGGAG
 ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
 AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAAGCACAGCATG
 GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTTCAGAATCCCAGAGACT
 CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
 TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
 GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
 GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAAGGAGGG
 AGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGATGGGGAGGGAGGGAAGGAAG
 GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAAAATAAATAAATTTTCAGGGATTATT
 ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
 GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
 ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT
 CCCACAAATGAAGAAAATGCTGTAAGAGTTTTTCACACACCGAAGTGGGAAACTTAA
 GGATTAGACAAGTCTAACAATGAGAATGGGGGAGAACAAAAAGAGACTGCACAGGG
 AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
 CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACAATAAAGTACATAT
 GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
 CACCCACTGTTTCACAGAGACTTTAAACATGAAACTGGAACCTATGTCTAGTGTTTTGA
 CTTAGGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
 CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTTCATGTATCCTGCATGAAATTC
 TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG
 AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
 TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
 GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTTCATGTTCTTT
 TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
 AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTT
 TTTGCTTCTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG
 ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACTGTGAGTGCTATGTAGAAACA
 GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTTCGGCTACATG
 ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
 AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
 TTAATAAAAAAAAAAACTATTGATGTGGAATGATCTTAATTTGGGGAAAAGACAATT
 TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
 ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAAATG
 AGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
 ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
 GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT
 GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATCTAA
 AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
 GGAAGTTGAAGTTTCATTGATAGTCTCATAAATCTTAAATGTATTTAACTGCTTTTTA
 CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTTCCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACCTATCTTGTTTGTGTTTCTTTTTTT
ATARKACGTATTTTCCTCAATTTCAATTTAGAATGATATCCCAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

Figure 51

MPKRAHWGALSVVLLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
 RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
 QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
 LSTIRHLYLAENMVRTLPAFMLRNMPLEENLYLQGNPWTCDCEMRWFLEWDAKSRGIL
 KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEEEQEQ
 EEDGGSQLEKLFQLPQWSISLNMDEHGNMVNLVCDIKKPMDEVYKIHNLQTDPPDIDIN
 ATVALDFECPMTRENYEKLWKLIAYYSEVPVKLHRELMLS KDPRVSYQYRQDADEEAL
 YYTGVRQAQILAEPEWVMQPSIDIQLNRRQSTAKKVLLSYYTQYSQTISTKDRQARGRS
 WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPDSKFSIL
 SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDVTVTIGKNPGES
 VTLPCNALAIPAEHLNRPNRRIINDLANTSHVYMLPNGTSLIPKVQVSDSGYYRCVAV
 NQQGADHFTVGITVTKKGSGLPSKRGRRPKAKALSRVREDIVEDEGGSGMGDEENTSR
 RLLHPKDQEVFLKTKDDAINGDKKAKKGRRLKLVKHSEKEPETNVAEGRRVFESRRR
 INMANKQINPERWADILAKVRGKNLPGTEVPPLIKTTSPPSLSLEVTPPFPAVSPPSASP
 VQTVTSAEESSADVPLLGEHHVLTGSSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
 SEKTEEITSTEGDLKGTAAPTLISEPYEPSPTLHTLDTVYEKPTHEETATEGWSAADVGSS
 PEPTSSEYEPPLDAVSLAESEPMQYFDPDLETKSQPDDEKMKEDTFAHLTPPTIWNDS
 STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKGMKEMSQTLQGGNM
 LEGDPTHRSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETTVGTLLDKDTTTVTTTPR
 QKVAPSSTMSTHPSRRRPNRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
 QVSSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGTTPRRKHGKRPKNHRYTPSTVSSRA
 SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDMYTTTRKIYSSYPKVQETLP
 VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
 GTPTWNPSTAQPGRLQTDIPVTTSGENLTDPLLKELEDVDFTSEFLSSLTVSTPFHQEE
 AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS
 TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTOHMSGPNE
 LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
 TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
 SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFTNKTLSTF
 PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPPLLHTPQTTGSPSTN
 LQNIPMVSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
 TAETDTVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
 QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTYVLGDTIAMECLAKGTPAPQI
 SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
 AALPPVIHQEKLENISLPPGLSIHHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
 NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTDVRYG
 GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVVKSVTDK

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVIFYGGDLKVDCVATGL
PNPEISWSLPDGSLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMRRVRVKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKTVWIHVNVQPPKING
NPNPITTVREIAAGGSRKLIDCKAEGIPTPRVLWAFPEGVVLPAPIYYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNCS
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNAA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPPGAGQGRFSWTLPNGMHLEG
PQTLGRVSLLDNGTLTVREASVFDRTYVCRMETEYGPSVTSIPVIVIAYPTRITSEPTPVI
YTRPGNTVKLNCMAMGIPKADITWELPKSHLKAGVQARLYGNRFLHPQGSLTIQHAT
QRDAGFYKCMANKILGSDSKTTYIHVF (SEQ ID NO: 21)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

[illegible][illegible]

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTCACGG
TGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCAG GTGCAAAGGC TCTTCCAGA GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCG GAAAAAGAAC CAGAGACCAA TGTTCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCGT TGATTAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTCTCTGCTG TTCTCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTAAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCA .
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAATACTACT TTTGCCTAGA
ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TTAACCCACACA TCAGATGGGAA AAGAAATTAA GGATGATGTT GCCACAAAATG
 TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
 GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
 AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
 GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
 GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
 TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCAAC
 AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
 GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
 CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
 CTGCTGCCCC GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
 ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
 AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
 TGGGGAATCC AGAAACAGAA GCAACCCCG TCAACAATGA AGGAACACAG
 CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
 ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
 GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
 GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
 AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
 TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
 TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
 ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATTC
 CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTAATCCAAA
 GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
 TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
 CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCAG
 ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
 TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
 CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
 ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
 CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
 CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
 GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
 TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
 TCGTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
 AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
 GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
 ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTACCCGT GCAGCAACCT
 CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
 CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTCCT
 GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CTGTGAGAGAGC CGCATCATACC TGCACGAAAA CCGGACCCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GCGGTCTATA AGTGCCTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTACAT TCACTGCACT GCCAAGGCTG CGCCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCCTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
GACGGGAGTC TGGTGAAGTC CTTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCCTCTCC TTATTCAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATT CAGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
 GGTTCTGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
 AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
 TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
 CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
 ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTTCTG
 GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
 CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
 ACATTCTCGG CAGTGACTCC AAAACAACCTT ACATCCACGT CTTCTGAAAT
 GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
 AGGGAAGCCA GGTGTTGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
 ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
 GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
 GAGACACTTT CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
 TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
 CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
 ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
 AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
 TCTTTCAGTT ATTCCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG
 GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
 TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTTCTTCC
 TGGAACTCAC TTTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
 ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAAAATTA
 ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
 ATATAATTTT AAAAAATTC TCTCCAACCT CCTTCAAATT CAGTCACCAC
 TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
 AGATTTCTT GTATGCAAAG TTTTGTGTTGA AAGCTGTGCT CAGAGGAGGT
 GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
 AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
 CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
 CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
 TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCCAAAA AAAAA

Human OCP: nucleotide sequence of ORF

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGC
CTGGTCGCCACCCCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCT
ACGGAGGTACACTGCACATTTTCGGTACCTGACTTCCATCCCAGACAGCATCCCGCCC
AATGTGGAACGCATCAATTTAGGATACAACAGCTTGGTTAGATTGATGGAAACAGAT
TTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGCATTACACA
ATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTAT
AATAAAGTCCGAAACTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGA
TTGCACATGGACCACAACAATATTGAGTTTATAAACCAGAGGTTTTTTATGGGCTC
AACTTTCTCCGCCTGGTGCCTTGGAAAGGAAATCAGCTCACTAAGCTCCACCCAGAT
ACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCATTAAAGTTCCTA
TACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTATATGCCT
GACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAG
TGGTTGTCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAAGATAGA
AGTCCCTCTAGTGCTCAGCAGTGTCCACTTTGCATGAACCCCTAGGACTTCTAAAGGC
AAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAACCATTGAC
TCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTCATC
TCTCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGATCAG
TCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCC
ATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTT
TTGGTGTGCAACATAGATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTG
TACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAG
CTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTACCAACATA
GAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCT
CAAATCACTTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATT
TCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTAGGTGGAACCGTTGGC
CTGAACTGCCCAGGCCAAGGAGACCCACCCACACGTGGATTGGCTTCTAGCTGAT
GGAAGTAAAGTGAGAGCCCCCTATGTCACTGAGGATGGACGGATCCTAATAGACAAA
AGTGGAATAATTGGAATCCAGATGGCTGATAGTTTTTGACACAGGCGTATATCACTGT
ATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAA
CCTTTGGTGAAGCCTATCAGGAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGCTGGGTT
ATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAACAAT
GGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTTGATTTTTTGGATTTTCCAAGTTTCAGTCAAGATGAAA
GGACAAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCC
AATCCTATTGCTCATCTTAAGGAGCCACCAGGTGCACAACCTCCGTACATCTGCTCTG
ATGGAGGCTGAGGTTGGAACACACCTCAAGCACAAAGTAAGAGGCACAACCTATCGG
GAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGG
AGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTGTTG
GAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCA
CCCCAGTGGTCACCCAACTCCCAAACATACTGGTGAAGAAGACGATTCTCAGGC
ATGCTCGCTCTACATGAGGAATTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCA
GCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATA
AATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAA
CCCACAGATTTCAAACGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATA

AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCATCCACTGTC
TTTCCACTGCTACTTGGAGCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGA
AGAGAGCATTTCCAAAGTAGACCCCCAATAACAGTAAGGACTATGATCAAAGATGTC
AATGTCAAAATGCTTAGTAGCACCACCAACAACTATTATTAGAGTCAGTAAATACC
ACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCACTTC
TATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACA
GCTGCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTT
AGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATAT
AGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTCAGGTCAACAACCAGAGGT
TCTTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACATGTCTG
TCCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCT
ACAACCTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAA
ACAACACCCACAATAAAATATTTTCAGGACTGAAATTTCCCAAGTGACTCCAACCTGGT
GCAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAAGTAAAC
GCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCAAGTGATTACA
TCGTCACTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGG
TTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAAACCAAAA
GGCAGATTAAGGAATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTT
CCTAAAACATCTCCTGCTTTACCACAGAGACAAAGTTCCCCTTTCCATTTACCACA
CTTTCAACAAGTGTGATGCAAATTCATCTAATACCTTGACTACCGCTCACCACACT
ACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTCCCA
CCCCTTAACCCTATGCTTCCCTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATC
ATATCAACGCAAACAGCAATACCAGCAACAACTCCTACCTTCCCTGCATCTGTCATC
ACTTATGAAACCCCAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAG
CCTCAAAAGAAGAACAGGACTGACCCAAACATCTCTCCAGACCAGAGTTCTGGCTTC
ACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACTCATTCCCCACCAGAA
AACACAACCTGGGATTTCAAGCACAAATCAGTTTTTCATTCAAGAACTCTTAATCTGACA
GATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAAATTGCT
TCTGAAACAACCTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCA
TTAGACACTCCCATACCACCTTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCC
ATCTCCCCCTCCCTTTACTCAGAGAGCAGTTACTGACACACGTGGCGACTCCCATTTC
CGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAATCTCCAA
ATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCATCT
CCCTGGGCAGAATACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAAGGC
AAAAAGCCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCTT
GTTTCAGATTGGGATGGACAGAAGAACAAGAAGAGTGACTTTGATAAGAAACCA
GTTCAAGAAGCAACAACCTTCCAAACTCCTTCCCTTTGACTCTTTGTCTAGGTATATA
TTTGAAGAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCTAAC
TCAGATGCCTTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCACCATTCAATTGG
ACCAGAGTTTCAGGACTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTT
CTCCCCAATGGTACCCTGTCCATCCAGAGGGTGGAATTCAGGACCGCGGACAGTAC
TTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGTCTGTG
GTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAGAGATCACAGTTCATTCC
GGAAGCACTGTGGAACCTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCT
GTGGTGACGGTTGACGGAACATTGGTCCCTCCACAATCTCAGTATTTATGACCGTGGC
TTTTACAAATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAATA
CAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGC
ACTTGGGGTGAAAGTTTAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGC

GTTTACTGGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCC
AAGTTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGAC
AGGGGCACTTATGAATGCATTGCTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTA
ATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGATAGAAGCTGCATCCCAGAAA
AGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAACTGCTCAGCCACTGGGGAG
CCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCGACCAGTGGAGC
TGGATCCACGCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGAC
AGTGGTGTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATG
CATGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAAG
CAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCCGGCTCCCCAGTG
CCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAACAATGCAATGCAAGCC
GATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACCTTTATAC
TTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGCTATGCCCAGAACACC
CTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATA
AGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCTTGAC
TGTGAGGTCCTACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTTCCAATGAC
ATGATTTTCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGGTCTTTGACCATC
AACAAAGTGAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAACTGGATGTGGTCTCTAAACCTCCATTAATCAAT
GGTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAA
CACTTTGACTGCAGAGCTGAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCA
GACAAATATTTTCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAAT
GGAACCTTGGAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTGTG
GCCCCGAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATG
CTGAGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGA
AAGTCCACAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGG
ATTTTACCAAATGGCACACGATTTTCCAATGGACCACAAAGTTATCAGTATCTGATA
GCAAGCAATGGTTCTTTTATCATTTCTAAAACAACCTCGGGAGGATGCAGGAAAATAT
CGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGA
GAATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGG
ACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATATTG
CATGACAATGGCACCTTAGTCATTAAGAAGCAACAGCTTATGACAGAGGAACTAT
ATCTGTAAGGCTCAAATAGTGTTGGTCATACACTGATTACTGTTCCAGTAATGATT
GTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACA
GGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACA
TGGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGA
AGTGAGCAGCTTCACTTACAAGGTACCCTAGTCATTGAGAATCCCCAAACCTCCGAT
TCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTATGCAGCAACG
TATATTCAAGTAATCTGA

Human OCP: predicted amino acid sequence

MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYMPTE VHCTFRYLTS
IPDSIPPNVE RINLGYNLSV RLMETDFSGL TKLELLMLHS NGIHTIPDKT
FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
PQEMVSYMPD LDSLYLHGPN WTCDCHLKWL SDWIQEKPDV IKCKKDRSPS
SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
SAFISPQGF M APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
YYKYKQVAPK PEDIFTNIEA DLRADPSWLM QDQISLQLNR TATTFSTLQI
QYSSDAQITL PRAEMRPVKH KWTMISRDN TKLEHTVLVG GTVGLNCPGQ
GDPTPHVDWL LADGSKVRAP YVSEDGRILI DKSGKLELQM ADSFDTGVYH
CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
PDASISWVIP GNNVLYQSSR DKKVLNNGTL RILQVTPKDQ GYYRCVAANP
SGVDLFIFQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
SALMEAEVGK HTSSTSKRHN YRELTQRRG DSTHRRFREN RRHFPPSARR
IDPQHWAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVNSQI
LPPEEPTDFK LSTAIKTTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
TEFQDSDQMG RGREHFQSRP PITVRTMIKD VNVKMLSST NKLLLESVNT
TNSHQTSVRE VSEPRHNHFY SHTTQILSTS TFPSDPHTAA HSQFPIPRNS
TVNIPLFRFF GRQRKIGGRG RIISPYRTPV LRRHRSIFR STTRGSSEKS
TTAFSATVLN VTCLSLPRE RLTTATAALS FPSAAPITFP KADIARVPSE
ESTTLVQNPL LLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTSI
PMEKTHKVNA SYPRVSSSTNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
RKIPWQQNFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
HFTTLSTSV M QIPSNTLTTA HHTTTKTHNP GSLPTKKELP FPPLNPMLPS
IISKDSSTKS IISTQTAIPA TTPTFPASVI TYETQTERS R AQTIQREQEP
QKKNRTDPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
RTLNLTDVIE ELAQASTQTL KSTIASSETL SSKSHQSTTT RKASLDTPIP
PFLSSSATLM PVPIPPFTQ RAVTDTRGDS HFRMTNTVV KLHESSRHLN
QMPSSQLEPL TSSTSLLHS TMPALTTVK SQNSKLTPSP WAEYQFWHKP
YSDIAEKGKK PEVSMLATTG LSEATTLVSD WDGQKNTKKS DFDKKPVQEA
TTSKLLPFDS LSRIFFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
IHWTRVSGLD LSRGNQNSRV QVLPNGTSLI QRVEIQDRGQ YLCSASNLFG
TDHLHVTLSV VSYPPRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
LANQTVVSES SQGSRQAVVT VDGTLLVLHNL SIYDRGFYKC VASNPGGQDS
LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYVWLSD
GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRV
MLTMEERVTS PRIEASQKR TEVNFGDKLL LNCSATGEPK PQIMWRLPSK
AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNKMGD DLILMHVSLR
LKPAKIDHKQ YFRKQVLHGK DFQVDCKASG SPVPEISWSL PDGTMINNAM
QADDSGHRTR RYTLFNNGTL YFNKVGVAEE GDYTCYAQNT LGKDEMKVHL
TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
FSIDRYTFHA NGSLTINKVK LLDGSEYVCV ARNPSGDDTK MYKLDVVSKP
PLINGLYTNR TVIKATAVRH SKKHFDCAE GTPSPEVMWI MPDNIFLTAP
YYGSRITVHK NGTLEIRNVR LSDSADFCV ARNEGGEVL VVQLEVLEML
RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPVI

LTYAPGTVKG ISGESLSLHC VSDGIPKPNI KWTMPSGYVV DRPQINGKYI
 LHDNGTLVIK EATAYDRGNY ICKAQNSVGH TLITVPVMIV AYPPRITNRP
 PRSIVTRTGA AFQLHCVALG VPKPEITWEM PDHSLSTAS KERTHGSEQL
 HLQGTQVIQN PQTSDSGIYK CTAKNPLGSD YAATYIQVI*

[illegible]